nature portfolio

# Peer Review File

# Rapid, systematic updating of movement by accumulated decision evidence

Corresponding Author: Dr Manuel Molano-Mazón

**This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.**

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
In this manuscript, the authors explore how the dynamics of decision-making affect the kinematics of response movement. They employed a 2-alternative forced choice task (previously developed and published) in both rats and humans, and they found that the dynamics of evidence accumulation affect the kinematics of response movements. They developed a computational model that maps the decision variable onto the movement kinematics at discrete time points. This model successfully reproduced the behavioral data.

This study, among others, offers compelling evidence that the decision-making process and movement responses are more intertwined than previously believed. Specifically, here, the results show that evidence accumulated during the decisional process systematically updates the kinematics of the response trajectory.

The paper addresses a timely question about the relationship between cognitive variables and movements (see for instance Zagha et al. 2022). The methodology is sound and, overall, the results support the conclusions and claims. I have only minor comments detailed below in order of priority.

1) Is there a behavioral signature of the discrete (2-readout) model that would argue against a continuous model? This could be, for example, evidence suggesting the presence of two distinct, stereotyped ballistic movements.

2) In Figure 1d, the mean response accuracy shows dependence on stimulus strength even at very short reaction times (including RT = 0). This finding is unexpected and not captured by the model (Fig. 6b). Could the authors elaborate on whether this might be attributable to trials where the rats adjusted their responses mid-decision (change in mind), or offer alternative explanations for this intriguing result?

3) Silent catch trials, where no auditory stimulus was played, were introduced to isolate the contribution of prior evidence on vigor. However, the manuscript lacks details on how behavior in these trials compared to non-silent ones. Did animals respond systematically, or were there missed trials? Were reaction times similar? Interestingly, the possibility of animals "changing their mind" during silent trials is also an open question.

4) While employing analogous tasks in rats and humans, the authors slightly modified the human version (introducing a response window and longer fixation time). Could they elaborate on the rationale behind these specific changes?

Reviewer #2

(Remarks to the Author)
This paper proposes an extension to classical evidence accumulation models, which now not just predicts choices but also explains full movement trajectories (extracted from freely moving rats' video data and human touchscreen swipes). The modelling makes a lot of sense, is an exciting extension of current accounts of evidence integration, and the paper is a pleasure to read. I applaud the authors on this nice piece of work, and look forward to seeing how these extended models will shape the field in the years to come.

major
1. Alternative models: alternative models are all very simple (and maybe not super plausible) - have you considered models that are a bit more interesting as alternatives (or hybrid, where parameters are not set to zero but rather allowed to be fit a weight)?
- the assumption is that there are just 2 times at which the decision variable is read out. why two (and not one every X seconds, perhaps in a theta rhythm that samples from working memory or so)? this is briefly mentioned in the discussion, but a bit more rationale may be nice when the model is presented.
- would a better alternative model (that would produce more realistic CoM proportion than 50/50 initial movement direction) be to make the initial movement in the direction of the prior, regardless of any evidence accumulated up until that point?
2. Please share your data and code. For such a novel model, providing others with clear instructions to fit this model would be a major boost to applying these ideas across more datasets in the future (e.g. integrate into an existing fitting toolbox like HDDM?)

minor
- I could not open the video, either downloaded from the Nature portal or from bioRxiv
- evolutionary strategies: I don't think we can conclude from this paper that the decision/movement strategy evolved 75 million years ago; it may have evolved in parallel afterwards, or it may be a cognitive consequence of some chaining of other, more primal evolutionary stages (see work by Paul Cisek). I'd remove this part.
- why not replicate the results one-to-one from rats to humans (e.g. include fiogure 2f,g and some others) and then in the model? It's not super clear which figures are considered the true 'diagnostics' of the model. Perhaps in the supplement, it would be possible to match figures side-by-side between species for easier comparison (and for identifying where the species differ).
- is the model not fit on the human data - and if not, why not? Can you give more info about human data, e.g. did human response patterns change depending on Repeating/Alternating block?
- please mention that you use only male rats (why?)
- terminology: what is a reaction vs movement? Intuitively, the two terms could have been switched around (movement is when the animal first moves, reaction is when the choice is confirmed). The terms movement initiation / movement duration may also make sense, and be less ambigious as to the underlying process?
- interpret afferent and efferent delay time (suppfig): 20ms is very small, is that physiologically plausible?
- silent catch trials - more info would be good about proportion of trials in which rats moves congruent vs incongruent to their prior (as it could be expected that they would move incongruently on only a small number of trials) how does this fit with change of minds depending on an initial snapshot of sensory information https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1009738
- I'd predict from the model that CoMs should happen earlier in trajectory than vigour updates, as anytime you hit the threshold vs waiting for last afferent info. Is this the case?

Reviewer #3

(Remarks to the Author)
Nature Comms Review 2024-03-08

Overall this is a solid piece of work nicely combining animal experiment, human experiment, and a (simple) generative model. The results together are mostly convincing in the authors' main conclusion, that rats and humans continue to process information during movement and that the incoming information can alter the ongoing movements. The topic is of general interest and this work adds to an ongoing literature in the field, supporting and bolstering prior ideas and findings in a new setting. Most of my criticisms are purely aimed at clarifying the manuscript as some of the points, if I understood them correctly, would benefit from being made more forthrightly.

Overall it would be nice (but not necessary) to see clearly listed, perhaps even in a table, the key findings with the supporting figures and results. That is, given the model contains many elements that appear somewhat ad hoc, while they are fitted to the data it is good to know which ones are more certain, which ones are less certain, and where is the confirmatory evidence. Especially as some of the comparison models seemed designed to fail (in particular, why would the initial choice be random when there is a clear prior? Unless the authors mean using a probability-matching strategy for "random"—but it appears they mean 50-50?)

Examples where I would like to see more explanation follow:

p.3 "a stimulus was played … until the animal poked out …" this is confusing as the word "until" implies the stimulus stopped when the animal begins its movement, but the point of the paper is about the animal integrating new stimulus information during the movement. More clarity is needed here as on the same page "rats process the stimulus … while moving away from the central port" – is this even during the absence of incoming stimulus so the processing and changes of mind all occur after the stimulus is over? If that is the case, then more emphasis of this point would be helpful.

Fig 2h: perhaps indicate that the numbers in the legend (0 0.25 05 1) refer to stim strength if this is the case?

Fig 2j: Some comment on why the prior is negatively correlated with position soon after stim onset and then positively correlated with position longer after stim onset would be helpful. One might expect the shift to be from high correlation with prior to less correlation with prior as stimulus information is integrated.

p.12-13 Model explanation is confusing, e.g. what determines “x1” and what determines the time of “threshold collapse” which seems directly related to the motor initiation time? The theta_DV threshold is not clearly defined – is it the same as “theta” on the graph in Fig 1A? But then it appears that the level “theta” (or -theta) is never reached because the thresholds collapse. Perhaps that is because an “Action Initiation” process has reached threshold, in which case for clarity it is essential that process is included in the figure (and I apologize if it already is!). If there are these two types of methods for initiating action (accumulation to theta_DV *or* action inititation) then I think both methods should be shown and explained clearly in the figure.

p.20: on the similarity of strategy suggesting commonality over 75 million years, this sentence ignores the (commonplace) occurrence of convergent evolution.

p. 24 The explanation of Action Initiation includes a “Trial Index” factor in the drift rate – does that mean there is a monotonic change over trials in likelihood of action initiation? “Trial index” is neither defined nor mentioned till later (Supp Fig 1), so worth adding something about it and why it is a factor in the model when first mentioned.

p.24 “until the stimulus is fully integrated”: this is confusing – what determines whether it is fully or only partially integrated? I would have thought the magnitude of x2 determines how much it has been integrated and this is inextricably tied in with whether it has reached a threshold or no, so the “or else” is confusing. Do you mean until “stimulus offset”? But in that case no clear offset time has been designated as far as I can tell.

p.26 Within Alternative Models it is surprising that the authors did not assume, instead of “random initial choice” rather “initial choice equivalent to the prior” as that seems less of a “straw man” model for comparison. In fact from the diagram shown in Fig. 5 with such a short time between onset of stimulus integration and the collapse of the decision bound it seems that the initial choice would always be based on the prior. So a bit of explanation is needed as to how often that does and does not happen i.e. how well the data is well-fitted by a model that simply randomly commences a movement in the direction indicated by the prior.

Minor Typos:
There are quite a few grammatical errors of English, particularly in the Introduction and Discussion, which could easily be corrected by sending the text through a GLM. e.g., in the first few pages:

p.2 “allows *us* … to” or one can say “allows for the prediction of …”
p.3 “Subjects … *their*”
p.3 “reverting *to* the”
p.6 “of a camera”: I assume taken “by a camera” and “of the animal”?
p.6 “consisting *of*”
p.6 “impact … on trajectories”

Reviewer #4

(Remarks to the Author)
Dear authors,

First, my sincere apologies for the late review. This is a topic close to my heart and as such I wanted to read it thoroughly but initially lacked time and focus.

This paper is a joy to read. You systematically combine research fields that have mainly been in parallel universes. I appreciate how you bring continuous (actually intermittent) updating of movement trajectories and especially vigor into the domain of decision making.

The experimental results seem crystal clear, although you do need a couple of data analysis tricks that require better explanation. I do understand how you introduced biases using P_rep of stimulus, but how you converted this into 'z' remains unclear to me (I did read methods and supplement but would need the earlier publication to really get it. Can you give a bit more of an intuition? - Pg 3 does not suffice with 'logistic regression'). So, I am very supportive to your study, but would like you to convince me a bit more with statistical tests and proper identifiability of the model. Also a deeper discussion about the pro's and con's of the model (it is rather descriptive in place) would appreciated.

Pg 3: The ref. to Hermoso-Mendizabal suggest that you might have re-used that dataset. I don't believe you did. Make clear that this ref. is about an analysis trick, not about this dataset.

Figure 1F and 1I, please indicate in the panels the difference. Now you can only infer with hard work from the caption the difference between the two.

Figure 1G: What is the quality (r2 for example) of these regressions? The is clearly a significant influence, but how much variance remains unexplained (in the supplement you do indicate that there is a trial effect that you also regress out).

Pg 2: Visser et al. 2023 JNP have also shown changes in the unfolding trajectory and looked into stretch reflexes. May go with references to Resulaj and Van den Berg - also in other places.

Pg 4: Nice description of finding with proper reference to the figures, but I miss any form of statistical evaluation! I know that we are moving away from P-values, but some form of statistical inference should at least be provided.

Figure 2D: how was movement onset defined? Might be hidden far in the supplement, but would be nice to already mention.

Splitting time calculation on page 8 (panel 2H): give at least a hint, detail can stay where they are.

Figure 3B and 3E: I guess you present in B the actual movements (including direction) and in E you have the chosen port always positive. Make clearer. Why are trajectories much smoother in E than in B?

Figure 3D: I don't understand this figure. Explain better.

Page 10 "We explicitly inserted an urgency component" -> did you pilot and not find anything like the rats without? If so, indicate.

Again, for the human subjects you report figures, but no statistics or comparisons.

First, I like the simplicity of the model, with an interaction between decision variable, urgency signals and movement trajectory. Only with the supplement in. hand was I really able to work out what Figure 5 describes, but I don't see how to make it simpler. Nevertheless, I would like you to be even more clear about the fact that your model is intermittent in selecting and updating trajectories. Also, the choice of MT is not based on a process, it assumed to be known (driven by 'visor'?). I know that most OFC models also do not have an emerging movement time, but they are at least not intermittent. Please discuss why you choose intermittent updating (for computational feasibility in part?). I like how you use sophisticated techniques to determine the parameters of the model. I am not sure this should go into Figure 5. It could be a methodological supplement.
I appreciate how you find the parameters and construe a number of checks on the fitting procedure. However, the Apprimate Likelihood in the supplement is not that convincing (Supplementary Figure 4, red vs blue). You might say that Figure 5 provides proof, but that is only the stability of finding the same solution over-and-over again. I would have liked to see a response dataset based on a known generative parameter set and then fit and compare the posterior set. Also, on the real data the model does not completely capture the kinematics (well that is also not expected given the simplicity of the simulated minimum jerk trajectories), especially in Figure S6 all experimental traces are later and more shallow then those from the model. Please comment (I still like the model very much, but want you to reflect on the aspects it does not capture well). In the supplement you also show density distributions for MT, RT and proportion choices (the latter only for the model, can you compare these to the actual ones). I would have liked it if you would for example shuffle the datase (link with simulated data from another animal) and show that you do much better than chance.

Finally Figure S10 can be done much better. Why not build (per rat) a corner plot of the parameter distribution. This would show the consistency of the findings. The current figure is rather uninformative. Also, from the efferent and afferent estimates it seems that they are discretely sampled. Is that true, if so mention. If not, why are so many point at the same value. Can you reflect why sigma_MT reached the upper bound for all animals? What does that tell you about the model?

Pg 21: Only 6 rats had silent trials. Indicate in main paper. These trials play a key role in your analysis, but for the majority of the rats you don't have them.

Pg 20: just before Making decisions in a continuous world: 'indicating a lack of bias', can you clarify?

Pg 20: Making decisions in a continuous world: Not sure this is the paragraph, but I would like a discussion about intermittent and continuous control adjustments. For example. there is a 2021 paper by Enachescu, Schrater, Schaal and Vassilios that combine a neurodynamic field model with OFC to model trajectory averaging when making a choice. I believe this should be discussed in relation to your findings.

Pg 21: Model of prior evidence:
Sorry, I cannot decipher this. Probably if I would know or study the Ermine-Mendizabal paper I would. Can you write this more clearly and maybe even provide a cartoon?

Pg 26, just before Alternative Models: 'avoid identifiabilty issues'. Sure, but the you should also show that your model parameters are identifiable in the current version of the model (see comments above).

Same paragraph, please explain all variables. I can deduce most, but am quite well versed in this domain. For a PhD-candidate this is hard, if not impossible, to decipher.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
The authors have effectively addressed all my previous comments and those of the other reviewers through insightful additional analyses. The newly incorporated information is particularly valuable, enhancing both the clarity of the manuscript and the impact of the presented results. I recommend this manuscript for publication in its current form.

Reviewer #2

(Remarks to the Author)

I thank the authors for their clarifications and updates to the work. I am a bit disappointed that they were not able to share the code to allow us to review this alongside the manuscript, but leave it to the discretion of the editor whether this is mandated at submission.

Minor: Perhaps indicate that you use 'RT' and 'MT' in a way that may be different with respect to other literature involving continuous rodent choices. For instance, https://doi.org/10.1101/2023.07.04.547681 uses the terms 'movement onset' to denote what you here call 'RT', namely the first movement made that does not yet indicate the animal's choice.

Reviewer #3

(Remarks to the Author)
In the revised manuscript, the authors have done an exceptionally thorough job of addressing the prior critiques, for which I applaud them. I do think that between the editorial office and the authors a final grammatical sweep is needed (ensuring prepositions are correct in several places) but otherwise I have no concerns and think this is a very nice article.

Reviewer #4

(Remarks to the Author)
I applaud the authors with this great study. All four reviewers had constructive and often similar comments, that were taken up in a constructive manner by the authors. The reply and paper are a joy to read. I am impressed how the authors build a continuous version of their model but clearly indicate that it is hard to disprove one or the other. I believe this paper is a great addition to the field and deserves publication.

**Open Access** This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.
In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.
The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this license, visit https://creativecommons.org/licenses/by/4.0/

# Intro

We thank the reviewers for their constructive feedback. In addressing their comments, the paper has gained width and depth. We have made an effort in explaining the points that were not clear in our first submission. Here are the main changes:

- We have modified the text to clarify the points raised by the reviewers.
- We made the following modifications in the main figures:
    - Fig. 1d: changed binning of reaction times.
    - Fig. 1e-f, i: added titles.
    - Fig. 1g, i: added information about statistics.
    - Fig. 2h: updated legend.
    - Fig. 3b, e: added titles.
    - Fig. 5a: added action initiation process.
- We made the following modifications in the supplementary figures:
    - We added to supp. fig. 3 a panel showing the psychometric curves for humans conditioned on repetition/alternation blocks (supp. fig. 3a) and the regression weights of MT as a function of prior, stimulus and trial index (supp. fig. 3d).
- We added the following supplementary figures:
    - S5: Model fit to human data.
    - S11: behavior in silent trials.
    - S12: alternative models.
    - S16: parameter recovery.
- We added a supplementary table 1, summarizing the main findings of the study.

Overall, we believe the manuscript has improved substantially. We address the reviewers' comments point by point below. Text in green contains our responses, text in blue corresponds to extracts from the manuscript that have been added to assist reviewers in following the discussion and text in dark blue corresponds to text already present in our first submission that we copied here to provide some context.

# Reviewer #1 (Remarks to the Author):

In this manuscript, the authors explore how the dynamics of decision-making affect the kinematics of response movement. They employed a 2-alternative forced choice task (previously developed and published) in both rats and humans, and they found that the dynamics of evidence accumulation affect the kinematics of response movements. They developed a computational model that maps the decision variable onto the movement kinematics at discrete time points. This model successfully reproduced the behavioral data.

This study, among others, offers compelling evidence that the decision-making process and movement responses are more intertwined than previously believed. Specifically, here, the

results show that evidence accumulated during the decisional process systematically updates the kinematics of the response trajectory.

The paper addresses a timely question about the relationship between cognitive variables and movements (see for instance **Zagha et al. 2022**). The methodology is sound and, overall, the results support the conclusions and claims. I have only minor comments detailed below in order of priority.

We thank the reviewer for her/his kind comments and thoughtful suggestions. We have added the reference to Zagha et al.

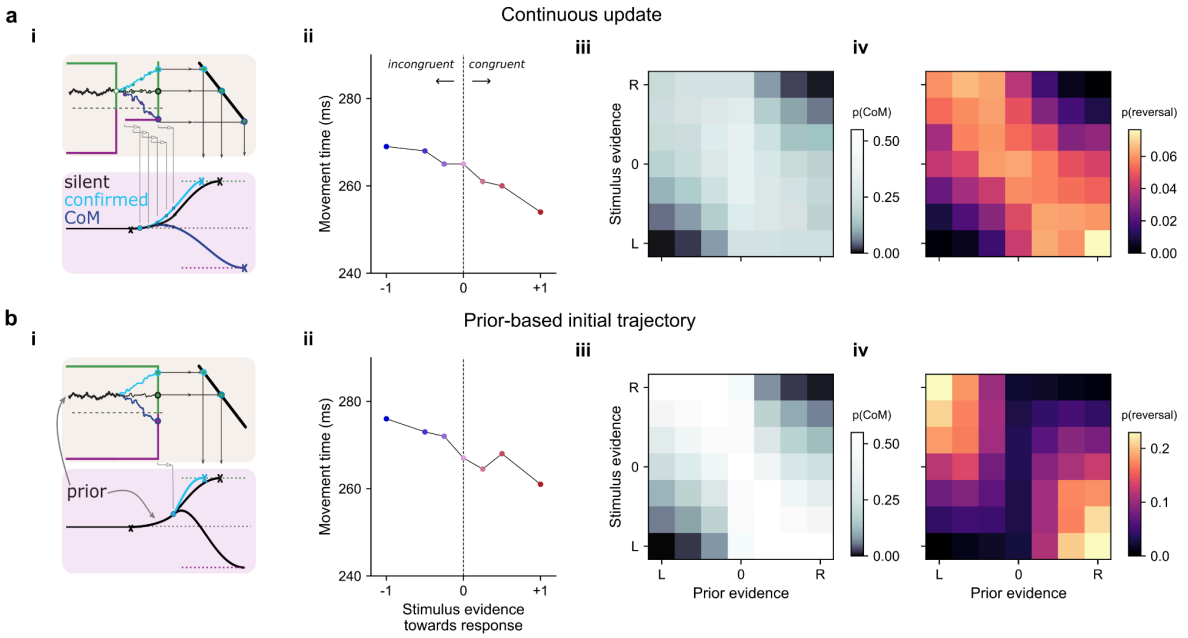
This new perspective allows us not only to predict motor features (the *how to do it*) from cognitive variables (that determine *what to do*), but also to access the cognitive processes underlying the production of a given behavior through the study of motor kinematics (Seideman, Stanford, and Salinas 2018; Korbisch et al. 2022; Zagha et al. 2022).

We addressed the rest of the suggestions point-by-point below.

1) Is there a behavioral signature of the discrete (2-readout) model that would argue against a continuous model? This could be, for example, evidence suggesting the presence of two distinct, stereotyped ballistic movements.

The question of whether the decision variable is read more than two times, or even continuously, is indeed interesting. As we discuss in the manuscript, there might be energetic constraints that impose a high cost on movements that faithfully follow the evidence accumulation dynamics including perturbations at a fast time scale where integration noise prevails. However, there are certainly other benefits to a more frequent readout, such as the ability to respond rapidly to sudden changes in the decision variable due to novel sensory information. One possibility for differentiating these two scenarios would be to assess whether novel sensory evidence is integrated into the trajectories with a fixed delay (as predicted by the continuous hypothesis) or only at two different time points (as predicted by the two-readout hypothesis). However, although in our paradigm sensory evidence fluctuated within each stimulus, animals typically relied on only one or two stimulus samples, which corresponds to the first 100ms, since the stimulus we used changes in magnitude every 50 ms (Hermoso-Mendizabal et al., 2020). A paradigm where stimulus evidence is resolved at a much finer time scale, such as the click task (Brunton et al., 2013 Science), would be more adequate to discriminate between the continuous and the 2-readout model. Another improvement in the experimental paradigm would be to increase the resolution of the camera since the sampling resolution of our camera is too poor (between 30 and 120 Hz, 65 Hz on average) to directly decompose fast components of the animal kinematics.

We have simulated a model that performs a continuous update of the trajectory based on DV and found that it captures the behavior of the animal about as well as the two-readout model. This has now been added as a supplementary figure (Supp. Fig. 12a):




**Supplementary figure 12.** Performance of alternative models. We evaluated the vigor and the trajectory reversals in two alternative models: a model where the vigor in the trajectory is continuously updated by reading out the evolution of the DV (instead of reading the DV in two discrete time points) (*Continuous update*, a) and a model where the initial trajectory only depends on the prior evidence instead of on the first readout of the decision variable (*Prior-based initial trajectory*, b). **i)** Cartoon illustrating each model. **ii)** Median movement time in express response trials (RT< 50 ms) with small prior (|z| < 10% percentile), as a function of stimulus evidence towards the selected port. **iii-iv)** Probability of CoM (iii) and reversals (iv) as a function of prior and stimulus evidence.

We have also modified the text in the results, discussion and methods sections to explain these new models.

Results

Note however that the experimental data did not allow to rule out alternative models where the initial trajectory is based on the prior alone (instead of the decision variable) or where the trajectory is continuously updated based on the decision variable (instead of a single update; (Supplementary Fig. 12).

Methods

Finally, in the *Continuous model,* after the first read-out is performed (as in the full model), the DV is read out every timestep, updating the movement time following Equation 2. As in the full model, the last read-out happens at $t_{aff}+t_{eff}$ from the first read-out or when the DV hits the change of mind bound.

Discussion (“Statistical modeling of response trajectories”)

One open question is whether more updates occur along the trajectory (Enachescu et al. 2021; Christopoulos, Bonaiuto, and Andersen 2015) or if the brain continuously transforms the decision variable into changes in the response trajectory (Friedman, Brown, and Finkbeiner 2013; Lepora and Pezzulo 2015). We found that a continuous-update model accounted for the experimental data equally well than our two-read-out model (Supplementary Fig. 12a). These two scenarios could be differentiated by assessing whether novel sensory evidence is integrated into the trajectories with a fixed delay (as predicted by the continuous hypothesis) or only at two different time points (as predicted by the two-readout hypothesis). However, although in our paradigm sensory evidence fluctuated within each stimulus, animals typically relied on only one or two stimulus samples (~100ms), and thus this analysis does not provide sufficient evidence towards either type of update scheme. A paradigm where stimulus evidence is resolved at a finer time scale, such as the click task (Brunton et al., 2013 Science), would be more adequate in that regard. Nevertheless, it is worth noting that the implementation of such a continuous model would require accommodating the dynamics of evidence accumulation, which typically exhibit fast sudden changes in the decision variable, to the constraints of optimal motor control which imposes smooth, gradual changes in velocities and acceleration.

2) In Figure 1d, the mean response accuracy shows dependence on stimulus strength even at very short reaction times (including RT = 0). This finding is unexpected and not captured by the model (Fig. 6b). Could the authors elaborate on whether this might be attributable to trials where the rats adjusted their responses mid-decision (change in mind), or offer alternative explanations for this intriguing result?

We appreciate Reviewer 1's insightful observation regarding Figure 1d. Indeed, we realized that the bins used for plotting the RT distributions were too large (10 ms), which precluded the difference between choices with short RTs (>5ms) and those with RTs approaching zero. We have now revised the panel accordingly, implementing bins that depend on RT as follows: bin = 5ms for RT<100ms and bin = 15ms for RT>100ms (to accommodate for the smaller RT density above 100 ms). We believe this provides a clearer depiction of the data. Here is how the updated panel looks:

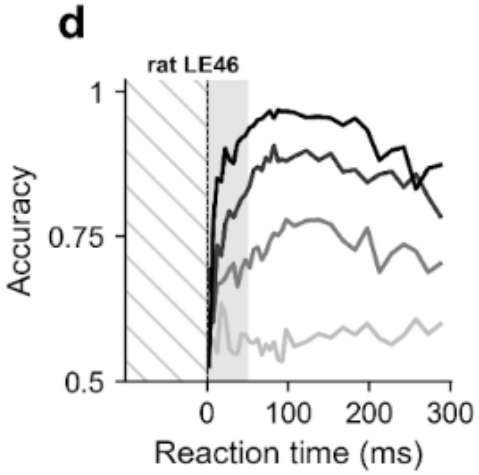
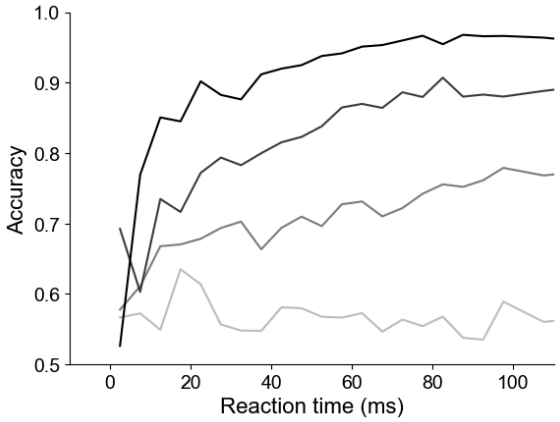


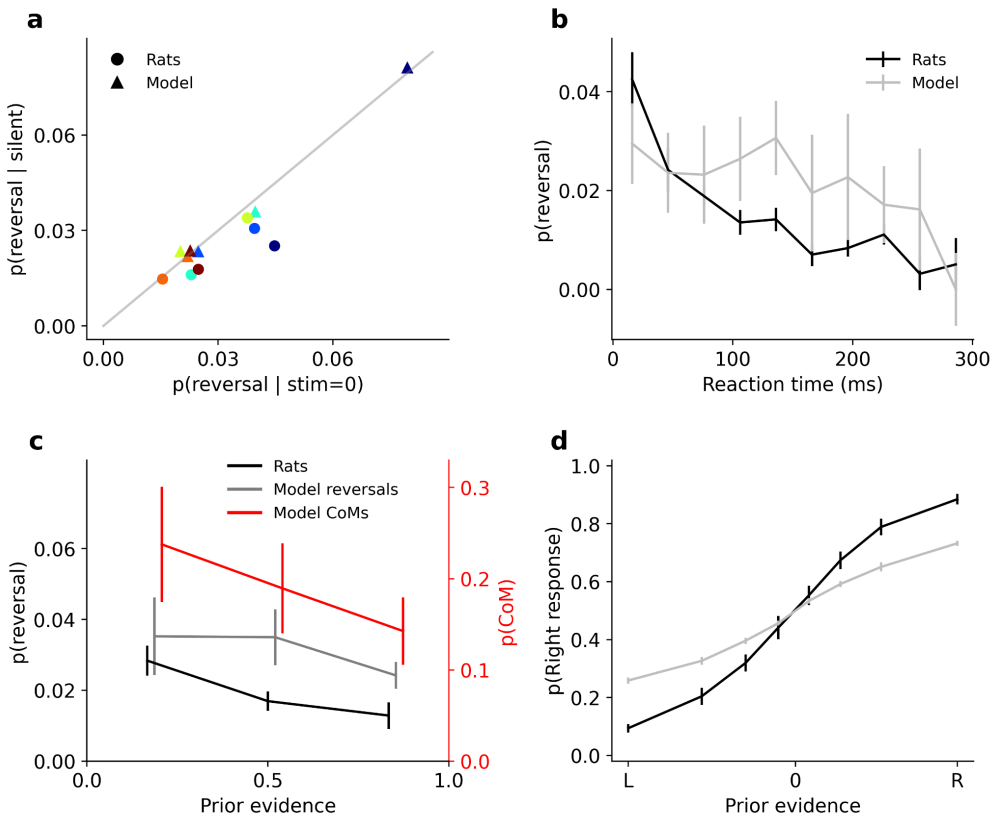
Figure. 1 [...] To accommodate for the smaller RT density above 100 ms, the bins were 5ms for RT<100ms and 15ms for RT>100ms.

We also copy here a zoom-in of the figure where the accuracy for short RTs can be appreciated in more detail:




3) Silent catch trials, where no auditory stimulus was played, were introduced to isolate the contribution of prior evidence on vigor. However, the manuscript lacks details on **how behavior in these trials compared to non-silent ones**. Did animals respond systematically, or were there missed trials? Were reaction times similar? Interestingly, the possibility of animals "changing their mind" during silent trials is also an open question.

We appreciate Reviewer 1's attention to the silent catch trials and their potential impact on isolating the contribution of prior evidence on vigor. We have now incorporated a supplementary figure (supp. fig. 11) showing the probability of reversals for silent VS trials with stimulus strength set at 0 (i.e. uninformative stimulus) (supp. fig. 11a), and their relationship to reaction time (supp. fig. 11b) and prior evidence towards the response (supp. fig. 11c), for both rats and fitted models. We found that reversals were actually equally frequent in silent trials than in these ambiguous non-silent trials, and that the proportion of reversals decreased with reaction times and prior towards the response. All phenomena are captured by the model. Additionally, we have added a panel showing the impact of prior evidence on choices in silent trials (supp. fig. 11d). All this evidence suggests that the behavior of the rats is very similar in silent and non-silent trials, with the only difference that the fluctuations in the decision variable are no longer driven by stimulus evidence.




**Supplementary figure 11.** Choices and reversals in silent trials, for experimental data and simulated model (n=6). **a)** Proportion of trajectory reversals in silent trials vs trials with stimulus strength 0. Each color corresponds to a different animal (circles: experimental data; triangles: simulated model). **b)** Proportion of trajectory reversals in silent trials as a function of reaction time, averaged across animals, for experimental data (black) and model simulations (gray). **c)** Proportion of reversals (black and gray) and CoMs (red) as a function of the prior evidence magnitude in silent trials. **d)** Proportion of rightward responses as a function of prior evidence in silent trials, for experimental data (black curve) and simulated model (gray curve).

This analysis is referred in the Results section:

To isolate the contribution of the prior evidence on vigor, we introduced, in a group of animals, *silent catch trials* where no acoustic stimulus was played (Group 2, n = 6 rats, mean±SD 6.7 ± 0.6% of all trials). In these trials, rats executed a normal response (median RT = 71.45 ± 16.37ms vs 70.30 ± 11.62ms, in silent vs non-silent trials, respectively), in which only the prior evidence could influence their final choice (see Supplementary Fig. 11d, black line).

[...]

Moreover, the model also accounts for the existence of trajectory reversals in silent trials (Supplementary Fig. 11a), even though silent trials were not used to fit the model. In our model, CoMs emerge in silent trials more often for weak prior evidence, when noise can more easily reverse the sign of the decision variable (Supplementary Fig. 11c). Finally, this

similarity between the model and experimental data also extended to human participants (Supplementary Fig. 5).

as well as in the Discussion:
Our model captures the dependence of trajectory reversals on stimulus, prior and reaction time, even though it was not fitted to any characteristics of the trajectory other than its overall duration. It also accounted for the unexpected presence of trajectory reversals in silent trials, hence illustrating how fluctuations in the decision variable can impact both choices as well as trajectories.

4) While employing analogous tasks in rats and humans, the authors slightly modified the human version (introducing a response window and longer fixation time). **Could they elaborate on the rationale behind these specific changes**?

The fixation time has been slightly increased in humans to make the overall pace of the experiment more comfortable for the subjects (which unlike rats did not benefit from extensive training on the task). The response window has been added to introduce an urgency component that forces the subjects to initiate their response based on prior expectations, matching rat intrinsic impatience. This is now more clearly explained in the methods section:

Subjects initiated each trial by maintaining the index finger of their preferred hand pressed against the starting point for 500 ms (fixation period). This adjustment of the fixation period was implemented because human subjects, unlike rats, were less accustomed to the task and required additional time to prepare for the trial. [...] Importantly, subjects had to start moving the finger within 300 ms after stimulus onset, otherwise the trial was considered a missed trial and the screen turned yellow. This time constraint introduced an urgency component which nudged participants into initiating their response early based on prior expectations (matching in this aspect rats' spontaneously impatient behavior).

## Reviewer #2 (Remarks to the Author):

This paper proposes an extension to classical evidence accumulation models, which now not just predicts choices but also explains full movement trajectories (extracted from freely moving rats' video data and human touchscreen swipes). The modelling makes a lot of sense, is an exciting extension of current accounts of evidence integration, and the paper is a pleasure to read. I applaud the authors on this nice piece of work, and look forward to seeing how these extended models will shape the field in the years to come.

We thank the reviewer for her/his kind comments and thoughtful suggestions, that we addressed point-by-point below.

### Major

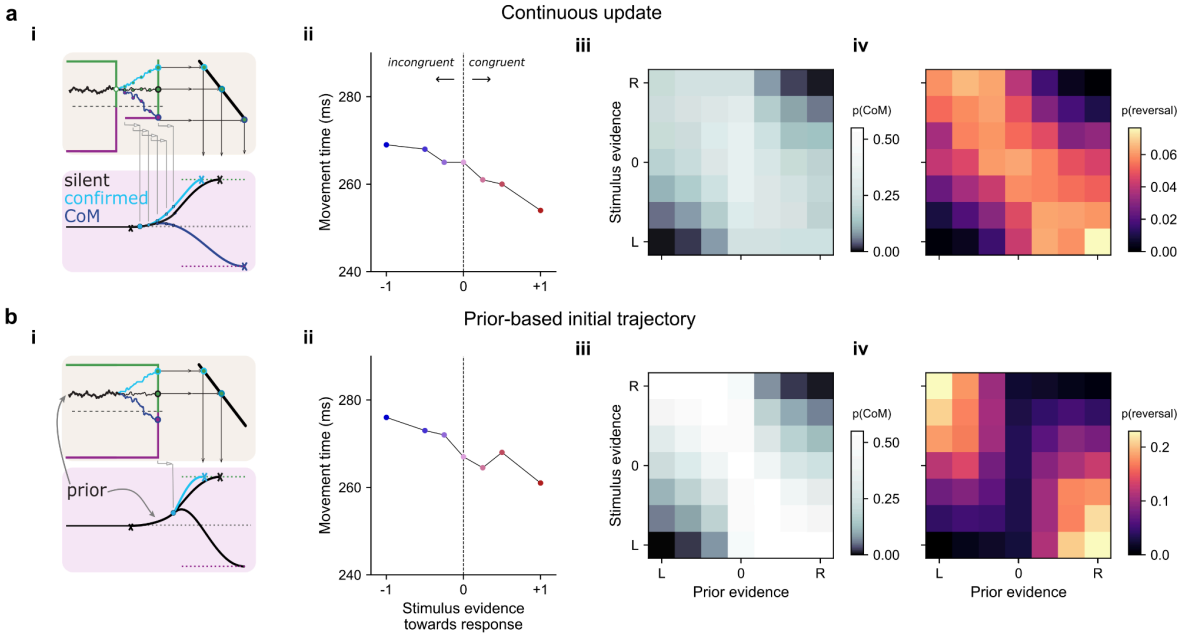
1) alternative models are all very simple (and maybe not super plausible) - have you **considered models that are a bit more interesting** as alternatives (or hybrid, where parameters are not set to zero but rather allowed to be fit a weight)?

2) the assumption is that there are just 2 times at which the decision variable is read out. **why two** (and not one every X seconds, perhaps in a theta rhythm that samples from working memory or so)? this is briefly mentioned in the discussion, but a bit more rationale may be nice when the model is presented.
3) - would a better alternative model (that would produce more realistic CoM proportion than 50/50 initial movement direction) **be to make the initial movement in the direction of the prior,** regardless of any evidence accumulated up until that point?

We thank Reviewer 2 for raising these important points. Indeed, the two alternative models we presented in Figure 7 are simple and aim to highlight the importance of three key aspects of our model: the informed nature of the initial trajectory, the existence of a second read-out and the late impact of the decision variable on vigor. Although we agree that it could be obvious that the model with random initial trajectory and the model with no trajectory update would fail to capture rats behavior, we have decided to keep them as they may help some readers. Nevertheless, we have now evaluated a supplementary figure with two new variants of the model with more interesting properties:

- in response to reviewer's point 2. As mentioned to reviewer 1, the question of whether the decision variable is read more than two times, or even continuously, is indeed interesting. As we discuss in the manuscript, there might be energetic constraints that impose a high cost on movements that faithfully follow the evidence accumulation dynamics including perturbations at a fast time scale where integration noise prevails. However, there are certainly other benefits to a more frequent readout, such as the ability to respond rapidly to sudden changes in the decision variable due for example to novel sensory information. One possibility for differentiating these two scenarios would be to assess whether novel sensory evidence is integrated into the trajectories with a fixed delay (as predicted by the continuous hypothesis) or only at two different time points (as predicted by the two-readout hypothesis). However, although in our paradigm sensory evidence fluctuated within each stimulus, animals typically relied on only one or two stimulus samples, which corresponds to the first 100ms, since the stimulus we used changes in magnitude every 50 ms (Hermoso-Mendizabal et al., 2020). A paradigm where stimulus evidence is resolved at a much finer time scale, such as the click task (Brunton et al., 2013 Science), would be more adequate to discriminate between the continuous and the 2-readout model. Another improvement in the experimental paradigm would be to increase the resolution of the camera since the sampling resolution of our camera is too poor (65 Hz on average) to directly decompose fast components of the animal kinematics.
- In response to point 3, we have now evaluated a model in which the initial readout is not random but is set by the prior (this was also suggested by Reviewers 3 & 4). This model yields results very similar to the original model. Although the experimental data does not allow to discriminate between the two models, we believe our model provides a more parsimonious explanation for the entire sensorimotor decision-making process, in which express responses coexist with trials featuring longer reaction times, where the system has time to integrate a significant amount of sensory evidence. In those trials, setting the initial response trajectory using only the prior when a more informed variable (the decision variable, integrating both prior and stimulus evidence) is available does not seem principled.

We copy below the supplementary figure 12 with these two new models and the sections that refer to these model in the Results and Discussion sections:




**Supplementary figure 12.** Performance of alternative models. We evaluated the vigor and the trajectory reversals in two alternative models: a model where the vigor in the trajectory is continuously updated by reading out the evolution of the DV (instead of reading the DV in two discrete time points) (*Continuous update*, a) and a model where the initial trajectory only depends on the prior evidence instead of on the first readout of the decision variable (*Prior-based initial trajectory*, b). **i)** Cartoon illustrating each model. **ii)** Median movement time in express response trials (RT< 50 ms) with small prior (|z| < 10% percentile), as a function of stimulus evidence towards the selected port. **iii-iv)** Probability of CoM (iii) and reversals (iv) as a function of prior and stimulus evidence.

## Results

Note however that the experimental data did not allow to rule out alternative models where the initial trajectory is based on the prior alone (instead of the decision variable) or where the trajectory is continuously updated based on the decision variable (instead of a single update; Supplementary Fig. 12).

## Methods

In the *prior-based initial trajectory model,* there is no reading of the decision variable in the first read-out. Thus, the initial choice is selected based on the sign of the prior (*z*) and the the initial movement time is set by its magnitude (i.e. $MT_{ini} = \beta_0 - \beta_{DV}|z_p z| + \beta_{TI} TrialIndex + \eta$; see Equation 1). The rest of the model is the same as in the full model. CoMs emerge when the first (set by *z*) and final choice are at odds. Finally, in the *Continuous model,* after the first read-out is performed (as in the full

model), the DV is read out every timestep, updating the movement time following Equation 2. As in the full model, the last read-out happens at $t_{aff}$+$t_{eff}$ from the first read-out or when the DV hits the change of mind bound.

Discussion (“Statistical modeling of response trajectories”)

One open question is whether more updates occur along the trajectory (Enachescu et al. 2021; Christopoulos, Bonaiuto, and Andersen 2015) or if the brain continuously transforms the decision variable into changes in the response trajectory (Friedman, Brown, and Finkbeiner 2013; Lepora and Pezzulo 2015). We found that a continuous-update model accounted for the experimental data equally well than our two-read-out model (Supplementary Fig. 12a). These two scenarios could be differentiated by assessing whether novel sensory evidence is integrated into the trajectories with a fixed delay (as predicted by the continuous hypothesis) or only at two different time points (as predicted by the two-readout hypothesis). However, although in our paradigm sensory evidence fluctuated within each stimulus, animals typically relied on only one or two stimulus samples (~100ms), and thus this analysis does not provide sufficient evidence towards either type of update scheme. A paradigm where stimulus evidence is resolved at a finer time scale, such as the click task (Brunton et al., 2013 Science), would be more adequate in that regard. Nevertheless, it is worth noting that the implementation of such a continuous model would require accommodating the dynamics of evidence accumulation, which typically exhibit fast sudden changes in the decision variable, to the constraints of optimal motor control which imposes smooth, gradual changes in velocities and acceleration.

[...]

Moreover, the position of the rats’ snout depended on the prior evidence already during fixation, before sensory integration started. We could thus confidently conclude that the initial trajectory is always informed by the prior. Our behavioral data however could not discriminate whether initial trajectories were based on the prior alone (as in the alternative model of Supplementary Fig. 12b) or also integrated the stimulus evidence available at that time, by reading out the decision variable (as in our main model). Indeed, the two models are essentially the same at short RTs, when the DV at the first readout only depends on prior evidence. However, they differ qualitatively at long RTs: while in the main model both CoMs and reversals readily decrease with RT (Fig. 6k and Supplementary Fig. 10c), in the alternative model the probability of CoMs increases with RT (as the mismatch between prior and DV increases). However those CoMs do not produce trajectory reversals as they occur immediately after the first read-out (because the DV is already strongly at odds with the initial prior-based-only decision). Thus, the two models make different predictions for the neural computations underlying the observed behavior. Apart from this difference, it should also be noted that a model where the initial trajectory and its update are based on read-outs of the same decision variable is more parsimonious.

**4) Please share your data and code.** For such a novel model, providing others with clear instructions to fit this model would be a major boost to applying these ideas across more datasets in the future (e.g. integrate into an existing fitting toolbox like HDDM?)

We agree with the reviewer, and plan to share the data and code upon publication. To ensure the replicability of our results, we will provide a comprehensive Python notebook. This notebook will demonstrate how to load, analyze, and model the data, as well as generate all figures presented in the paper. Furthermore, we emphasize that our approach can be adapted to fit the choice, reaction time (RT), and movement time (MT) of any dataset **without the need for extracting response trajectories**. This versatility could significantly improve parameter estimation of the drift-diffusion model (DDM) compared to traditional fitting methods based only on the pattern of choices and RTs.

## Minor

5) I could not open the video, either downloaded from the Nature portal or from bioRxiv.

We apologize for this. We transformed the video to mp4 format, which should be readable in all operative systems.

6) evolutionary strategies: I don't think we can conclude from this paper that the decision/movement strategy **evolved 75 million years ago**; it may have evolved in parallel afterwards, or it may be a cognitive consequence of some chaining of other, more primal evolutionary stages (see work by Paul Cisek). I'd remove this part.

We thank the reviewer for his/her comment. While we acknowledge that the conclusion drawn regarding the timing of the evolution of decision/movement strategies may be speculative, we believe that comparing behaviors across species can offer valuable insights into the underlying principles of sensorimotor processing. Therefore, rather than removing the paragraph, we have revised it to acknowledge the possibility of convergent evolution and the limitations in determining the timing of the development of these strategies. Here's the modified paragraph:

Making decisions in a continuous world

The behavior of rats and humans was very similar, despite performing very different types of motor response (full-body orienting vs. arm movements). This resemblance suggests that the ability to adjust movement endpoints and vigor while processing asynchronous information is a fundamental skill in the sensorimotor repertoire of vertebrates. At this point we can only speculate whether this skill evolved before the two species diverged around 75 million years ago (Striedter and Glenn Northcutt 2019) or if it arose independently. To this respect, it is worth noting that the basal ganglia, which regulates motor vigor (Turner and Desmurget 2010), is a highly conserved area since its emergence in the first vertebrates 500 million years ago (Grillner and Robertson 2016). Beyond their similarities, humans differed from rats as they performed many more reversals. Part of this difference may be due to the lower energetic cost of a change of trajectory in the finger compared to full-body responses.

7) why not replicate the results one-to-one from rats to humans (e.g. include figure 2f,g and some others) and then in the model? It's not super clear which figures are considered the true 'diagnostics' of the model. Perhaps in the supplement, it would be possible to match figures side-by-side between species for easier comparison (and for identifying where the species differ).

Spurred by this comment, we realized that the link between our main behavioral observations and the different elements of the model was not synthesized in the manuscript. We have thus included a table summarizing how each property of the model is backed by a specific behavioral signature found in rats in humans (Table 1). We thank the reviewer for his/her suggestion of adding a supplementary figure, but we believe the table provides a more informative and concise summary, serving the purpose effectively. Moreover, we think that the table could be used by readers for 'diagnostics' of their own behavioral dataset.

| **Properties of the model architecture** | **Behavioral signature** | Figure (rats) | Figure (humans) |
|---|---|---|---|
| Decision variable integrates stimulus and prior evidence | Choices depend on stimulus & prior evidence | Fig. 1b | Fig. 4c |
| There is an urgency signal that can trigger response movement onset. | In express responses, RTs do not depend on stimulus strength, so they must be triggered by a process other than the DV hitting a decision bound | Fig. 1c | not shown |
| Initial trajectory is determined by the first readout $x_1$ which is always influenced by prior evidence | Response trajectories correlate with prior evidence starting before movement onset. | Fig. 2j right | not shown |
| Movement vigor in the initial trajectory is modulated by the first readout $x_1$ | MT and velocity depend on prior evidence | Fig. 1d, f, 2d, f | Fig. 4 e |
| The initial trajectory is updated using the second readout $x_2$ , incorporating stimulus evidence | Trajectories split based on stimulus after some delay.<br>MT and velocity depend on stimulus evidence. | Fig. 2h-i<br>Fig. 2e, g | Fig. 4h<br>Fig. 4f |
| Movement vigor is updated based on the strength of the evidence at second readout relative to the first readout | MT in express responses with congruent prior are larger/smaller if the stimulus is congruent / incongruent, compared to silent trials | Fig. 1i | not shown |
| Changes of mind can occur after movement is initiated if the decision | Trajectory reversals emerge mostly when a trajectory | Fig 3c, i | Fig 4l |

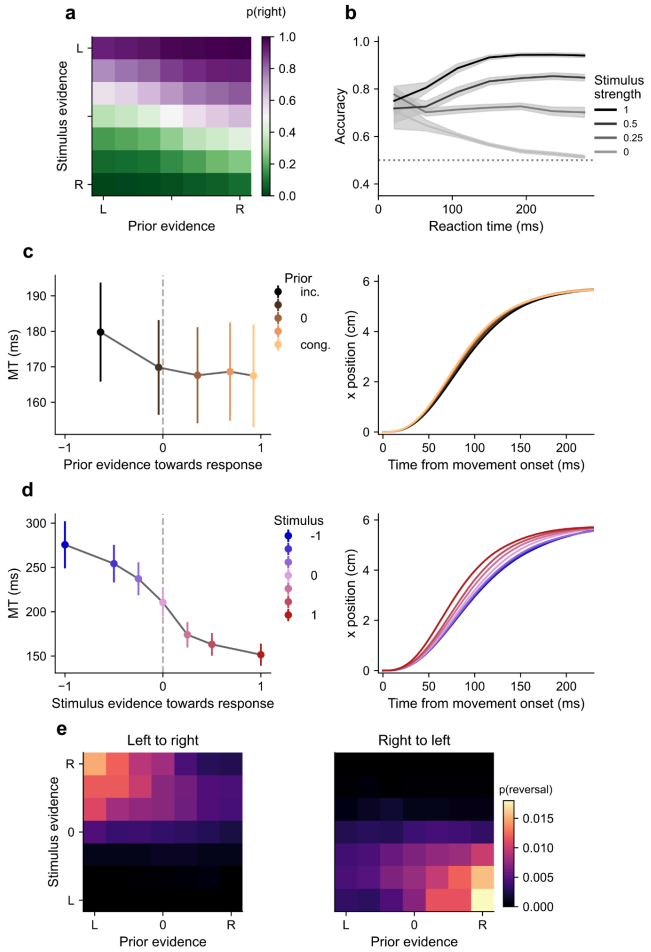
| | | | |
|---|---|---|---|
| variable reaches a certain boundary | guided by the prior is reverted by the stimulus | | |

**Table 1. Results supporting model architecture.** Each row explains the link between a property of the model (left column) and the behavioral signature that supports it (middle column). The right columns indicate the corresponding figures for the rat and human experiments.

The table is cited during the presentation of the model:
Although relatively complex, most of the properties of the model are derived from our previous observations (see Table 1 for an explicit link between observations and model features).

8) is the model not fit on the human data - and if not, why not?

This is a nice suggestion. We have now also fitted the model to the human data. Although the number of trials is much smaller in humans than in rats, the fitted model captured very well the behavior of human participants. We have added a new Supplementary Figure 5 summarizing the results from model fits on human data:



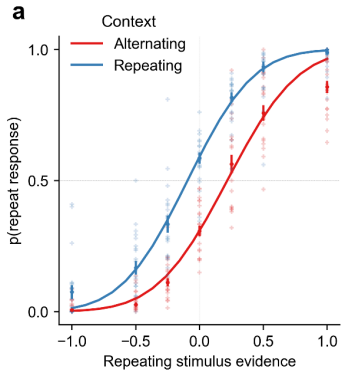

**Supplementary figure 5.** Simulations from the model fitted to human participants replicate the key features of human behavior. **a)** Proportion of right responses as a function of the stimulus and prior evidence (as in Figure 4c). **b)** Accuracy of the model as a function of RT, for different stimulus strengths (as in Figure 4d). **c)** Left: Median movement time as a function of prior evidence towards the

selected port. Right: Model trajectories depending on prior congruence (as in Figure 4e). **d)** Left: Average movement time across trials as a function of stimulus evidence towards the selected port (as in Figure 4f). Right: Average trajectory as a function of stimulus evidence, in low prior trials. **e)** Probability of reversals as a function of prior and stimulus evidence (as in Figure 4l). Error bars: standard error of the mean (s.e.m.). Note that the number of trials in humans is two orders of magnitude smaller than in rats, which hinders the approximation of the likelihood. Nevertheless the fitted models captured very well the behavior of human participants.

These new set of results reported in the Results section:
Finally, this similarity between the model and experimental data also extended to human participants (Supplementary Fig. 5). Overall, there is a compelling fit between our model and the experimental data, in terms of movement time, the timing of the impact of prior and sensory evidence onto trajectories, and the trajectory reversals.

9) Can you give more info about human data, e.g. did human response patterns change depending on Repeating/Alternating block?

We have now included a supplementary panel (panel a in supp. fig. 3) with psychometric curves showing the probability of repeat during repeating and alternating contexts. The figure shows that, as in rats (Hermoso-Mendizabal et al. 2020), human subjects tend to repeat/alternate more often in the repeating/alternating contexts.




**Supplementary figure 3. a)** Psychometric curves showing the probability of repeating the previous choice as a function of the stimulus evidence supporting a repetition, separately for alternating and repeating blocks, averaged over n=14 human participants.

10) please mention that you use **only male rats** (why?)

We used only male rats primarily due to space constraints within our animal facility. However, we acknowledge that including both female and male rats would have enabled us to draw more comprehensive conclusions. In ongoing projects, we are transitioning towards a balanced configuration to ensure a more representative sample in our studies. We now mention rats sex in the results:

We investigated how response trajectories are formed as male rats (Groups 1 and 2; n=15; see Methods)

11) terminology: what is a reaction vs movement? Intuitively, the two terms could have been switched around (movement is when the animal first moves, reaction is when the choice is confirmed). The terms movement initiation / movement duration may also make sense, and be less ambiguous as to the underlying process?

We thank the reviewer for this question as it provides an opportunity to clarify the terminology used in our paper. We define as “response movement onset” the moment at which the animal leaves the central port. This is now explicitly said in the main text:

In our experimental paradigm, the response movement onset (defined as the time at which the rat leaves the central port), was often driven by a proactive process, leading to a large fraction of reaction times (RTs)

Throughout the study, we used "reaction time" as the duration between stimulus onset and movement onset. Reaction time thus reflects the time taken for the subject to react to the stimulus. This definition is consistent with standard practices and is explicitly stated in the caption of Figure 1.

RT (reaction time): time from stimulus onset to initiation of rat response (central port withdrawal)

Conversely, we referred to "movement time" as the time taken to move from the central port to one of the lateral ports.

MT (movement time): time from response initiation to the lateral port entering.

Note that many previous experimental paradigms have conflated RT and MT into a single measurement. By separating them, we can investigate the differential impact of the evidence accumulation process on these distinct components of the response.

12) interpret afferent and efferent delay time (supp fig): 20ms is very small, is that physiologically plausible?

We thank the reviewer for this very relevant comment. These delays correspond to the minimal time for updating an existing trajectory, which have been shown to be smaller than those for initiating a new trajectory (Smeets et al. 2016 Motor Control). However, we acknowledge that the model fitting yielded quite implausible values for these delays. To address this, we have imposed that the sum of afferent and efferent delays ($t_{aff}$ + $t_{eff}$) is larger than 40ms, which is just below the minimum splitting time we have observed in our data (60ms). With this constraint, model fits provide values for $t_{aff}$ and $t_{eff}$ that are more in line with the experimental data. Furthermore, we have included additional text in the methods (“Model Fitting" section), acknowledging the discrepancy and potential limitations in fitting these parameters (see also reply to Reviewer 4):

Our model predicts afferent ($t_{aff}$) and efferent ($t_{eff}$) delays that appear unusually short. This discrepancy might be caused by the model failing to capture all the nuances of the behavior of the rats, e.g. some form of underfitting.

The change in the fitting boundaries of $t_{aff}$ and $t_{eff}$ did not alter the main results and claims of the paper. If anything, it actually improved the fit of the model to the experimental data. For example, the fitted model no longer produces reversals for the four rats that display an atypical pattern of reversals that do not correspond to prior-based initial trajectories that are corrected by the stimulus (Rats 36, 41, 85, 86 in Supp. Fig 2 and Supp Fig. 9). We updated the text accordingly in the results section:

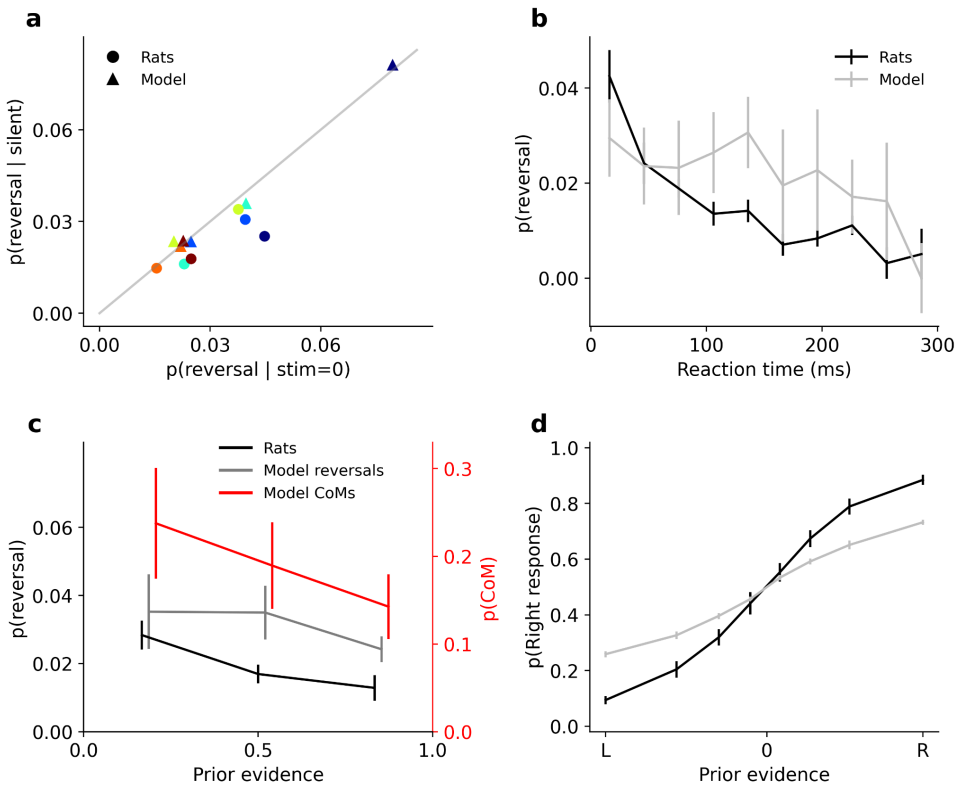
Our model predicts that only around a sixth of the CoMs lead to detectable trajectory reversals (CoMs represented 11.02% of the total, vs 1.98% for detected reversals).

and provide more information about the reversals produced by the model in the caption of Supp. Fig 9:

The fits in four rats (86, 41, 85 and 36) did not produce reversals. Notice however, that these animals displayed a nonstandard pattern of probability of reversals, that does not match a prior-based initial trajectory that is corrected by the stimulus (see Supplementary Fig. 2). Hence it is expected that the model does not capture this type of reversals.

13) silent catch trials - more info would be good about proportion of trials in which rats moves congruent vs incongruent to their prior (as it could be expected that they would move incongruently on only a small number of trials) how does this fit with change of minds depending on an initial snapshot of sensory information

We have now added a supplementary figure showing the behavior of the rats during silent trials. In particular, we show the impact of prior evidence on choices in silent trials (supp. fig. 11d). Additionally, we have incorporated three other panels (supp. fig. 11a-c) showing the probability of reversals for silent VS zero-coherent trials, and probability of reversals as a function of reaction time and prior evidence towards the response, for both rats and the model. The panels show that reversals were actually equally frequent in silent trials than in ambiguous non-silent trials, that the proportion of reversals decreased with reaction times and with the prior evidence towards the response. All phenomena are captured by the model. All this evidence suggests that the behavior of the rats is very similar in silent and non-silent trials.




**Supplementary figure 11.** Choices and reversals in silent trials, for experimental data and simulated model (n=6). **a)** Proportion of trajectory reversals in silent trials vs trials with stimulus strength 0. Each color corresponds to a different animal (circles: experimental data; triangles: simulated model). **b)** Proportion of trajectory reversals in silent trials as a function of reaction time, averaged across animals, for experimental data (black) and model simulations (gray). **c)** Proportion of reversals (black and gray) and CoMs (red) as a function of the prior evidence magnitude in silent trials. **d)** Proportion of rightward responses as a function of prior evidence in silent trials, for experimental data (black curve) and simulated model (gray curve).

- I'd predict from the model that CoMs should happen earlier in trajectory than vigor updates, as anytime you hit the threshold vs waiting for last afferent info. Is this the case?

Indeed, in the model, CoMs happen earlier than vigor updates. However, we could not find any reliable test to assess this in the experimental data. In principle, we would predict an earlier splitting time for trials when CoMs are frequent (e.g. opposite prior and stimulus) versus trials with infrequent CoMs (e.g. prior & stimulus evidence in the same direction). However, the deflections in CoMs are much larger than in confirmatory updates, which could lead to artifactual detection of earlier splitting time in trials with frequent CoMs.

# Reviewer #3 (Remarks to the Author):

Nature Comms Review 2024-03-08

Overall this is a solid piece of work nicely combining animal experiment, human experiment, and a (simple) generative model. The results together are mostly convincing in the authors'

main conclusion, that rats and humans continue to process information during movement and that the incoming information can alter the ongoing movements. The topic is of general interest and this work adds to an ongoing literature in the field, supporting and bolstering prior ideas and findings in a new setting. **Most of my criticisms are purely aimed at clarifying the manuscript as some of the points**, if I understood them correctly, would benefit from being made more forthrightly.

We thank the reviewer for her/his kind comments and thoughtful suggestions, that we addressed point-by-point below.

1) Overall it would be nice (but not necessary) to see clearly listed, perhaps even in a table, the key findings with the supporting figures and results. That is, given the model contains many elements that appear somewhat ad hoc, while they are fitted to the data it is good to know which ones are more certain, which ones are less certain, and where is the confirmatory evidence. Especially as some of the comparison models seemed designed to fail (in particular, why would the **initial choice be random when there is a clear prior?** Unless the authors mean using a probability-matching strategy for "random"—but it appears they mean 50-50?)

This is a really nice suggestion. We have now added a supplementary table that links the different elements in the model with specific findings supporting each of them. We hope that, through this table, readers can better understand the rationale behind different elements of the model and the extent to which they are supported by empirical evidence.

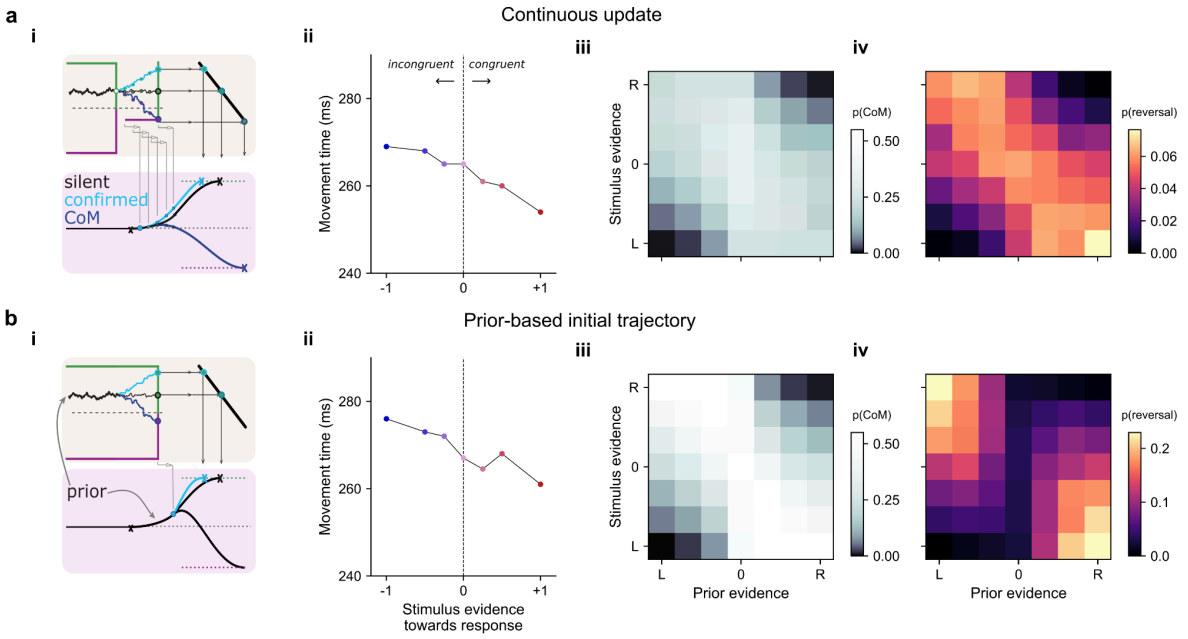
| **Properties of the model architecture** | **Behavioral signature** | Figure (rats) | Figure (humans) |
|---|---|---|---|
| Decision variable integrates stimulus and prior evidence | Choices depend on stimulus & prior evidence | Fig. 1b | Fig. 4c |
| There is an urgency signal that can trigger response movement onset. | In express responses, RTs do not depend on stimulus strength, so they must be triggered by a process other than the DV hitting a decision bound | Fig. 1c | not shown |
| Initial trajectory is determined by the first readout $x_1$ which is always influenced by prior evidence | Response trajectories correlate with prior evidence starting before movement onset. | Fig. 2j right | not shown |
| Movement vigor in the initial trajectory is modulated by the first readout $x_1$ | MT and velocity depend on prior evidence | Fig. 1d, f, 2d, f | Fig. 4 e |
| The initial trajectory is updated using the second readout $x_2$ , incorporating stimulus | Trajectories split based on stimulus after some delay. | Fig. 2h-i Fig. 2e, g | Fig. 4h Fig. 4f |

| | | | |
|---|---|---|---|
| evidence | MT and velocity depend on stimulus evidence. | | |
| Movement vigor is updated based on the strength of the evidence at second readout relative to the first readout | MT in express responses with congruent prior are larger/smaller if the stimulus is congruent / incongruent, compared to silent trials | Fig. 1i | not shown |
| Changes of mind can occur after movement is initiated if the decision variable reaches a certain boundary | Trajectory reversals emerge mostly when a trajectory guided by the prior is reverted by the stimulus | Fig 3c, i | Fig 4l |

**Table 1. Results supporting model architecture.** Each row explains the link between a property of the model (left column) and the behavioral signature that supports it (middle column). The right columns indicate the corresponding figures for the rat and human experiments.

The table is cited during the presentation of the model:
Although relatively complex, most of the properties of the model are derived from our previous observations (see Table 1 for an explicit link between observations and model features).

Regarding the alternative model with random initial choice, although we agree that this model should be expected to fail, we are not sure how obvious this will be for some readers. Nevertheless, following the reviewer's suggestion, we have now added a supplementary figure (supp. fig. 12b) with a model in which the initial trajectory is set by the prior. The model yields results very similar to the original model. However, although the experimental data does not allow to discriminate between the two models, we believe our model provides a more parsimonious explanation for the entire sensorimotor decision-making process, in which express responses coexist with trials featuring longer reaction times, where the system has time to integrate a significant amount of sensory evidence. In those trials, it seems unnatural to set the initial response trajectory using only the prior when a more informed variable (the decision variable, integrating both prior and stimulus evidence) is available. Furthermore, we have assessed another alternative model in which there is a continuous readout of the decision variable, instead of simply two read-outs. We copy below the new supplementary figure as well as corresponding text in Results and Discussion:




**Supplementary figure 12.** Performance of alternative models. We evaluated the vigor and the trajectory reversals in two alternative models: a model where the vigor in the trajectory is continuously updated by reading out the evolution of the DV (instead of reading the DV in two discrete time points) (*Continuous update*, a) and a model where the initial trajectory only depends on the prior evidence instead of on the first readout of the decision variable (*Prior-based initial trajectory*, b). **i)** Cartoon illustrating each model. **ii)** Median movement time in express response trials (RT< 50 ms) with small prior (|z| < 10% percentile), as a function of stimulus evidence towards the selected port. **iii-iv)** Probability of CoM (iii) and reversals (iv) as a function of prior and stimulus evidence.

## Results

Note however that the experimental data did not allow to rule out alternative models where the initial trajectory is based on the prior alone (instead of the decision variable) or where the trajectory is continuously updated based on the decision variable (instead of a single update; Supplementary Fig. 12).

## Methods

In the *prior-based initial trajectory model,* there is no reading of the decision variable in the first read-out. Thus, the initial choice is selected based on the sign of the prior (*z*) and the the initial movement time is set by its magnitude (i.e. $MT_{ini} = \beta_0 - \beta_{DV}|z_p z| + \beta_{TI} TrialIndex + \eta$; see Equation 1). The rest of the model is the same as in the full model. CoMs emerge when the first (set by *z*) and final choice are at odds. Finally, in the *Continuous model,* after the first read-out is performed (as in the full model), the DV is read out every timestep, updating the movement time following Equation 2. As in the full model, the last read-out happens at $t_{aff}$+$t_{eff}$ from the first read-out or when the DV hits the change of mind bound.

Discussion (“Statistical modeling of response trajectories”)

One open question is whether more updates occur along the trajectory (Enachescu et al. 2021; Christopoulos, Bonaiuto, and Andersen 2015) or if the brain continuously transforms the decision variable into changes in the response trajectory (Friedman, Brown, and Finkbeiner 2013; Lepora and Pezzulo 2015). We found that a continuous-update model accounted for the experimental data equally well than our two-read-out model (Supplementary Fig. 12a). These two scenarios could be differentiated by assessing whether novel sensory evidence is integrated into the trajectories with a fixed delay (as predicted by the continuous hypothesis) or only at two different time points (as predicted by the two-readout hypothesis). However, although in our paradigm sensory evidence fluctuated within each stimulus, animals typically relied on only one or two stimulus samples (~100ms), and thus this analysis does not provide sufficient evidence towards either type of update scheme. A paradigm where stimulus evidence is resolved at a finer time scale, such as the click task (Brunton et al., 2013 Science), would be more adequate in that regard. Nevertheless, it is worth noting that the implementation of such a continuous model would require accommodating the dynamics of evidence accumulation, which typically exhibit fast sudden changes in the decision variable, to the constraints of optimal motor control which imposes smooth, gradual changes in velocities and acceleration.

[...]

Moreover, the position of the rats’ snout depended on the prior evidence already during fixation, before sensory integration started. We could thus confidently conclude that the initial trajectory is always informed by the prior. Our behavioral data however could not discriminate whether initial trajectories were based on the prior alone (as in the alternative model of Supplementary Fig. 12b) or also integrated the stimulus evidence available at that time, by reading out the decision variable (as in our main model). Indeed, the two models are essentially the same at short RTs, when the DV at the first readout only depends on prior evidence. However, they differ qualitatively at long RTs: while in the main model both CoMs and reversals readily decrease with RT (Fig. 6k and Supplementary Fig. 10c), in the alternative model the probability of CoMs increases with RT (as the mismatch between prior and DV increases). However those CoMs do not produce trajectory reversals as they occur immediately after the first read-out (because the DV is already strongly at odds with the initial prior-based-only decision). Thus, the two models make different predictions for the neural computations underlying the observed behavior. Apart from this difference, it should also be noted that a model where the initial trajectory and its update are based on read-outs of the same decision variable is more parsimonious.

Examples where I would like to see more explanation follow:

2) p.3 “a stimulus was played … until the animal poked out …” this is confusing as the word “until” implies the stimulus stopped when the animal begins its movement, but the point of the paper is about the animal integrating new stimulus information during the movement. More clarity is needed here as on the same page “rats process the stimulus … while moving

away from the central port" – is this even during the absence of incoming stimulus so the processing and changes of mind all occur after the stimulus is over? If that is the case, then more emphasis of this point would be helpful.

We thank the reviewer for this question, as it draws attention to the need for clarity regarding the integration of stimulus information during the rats' movement. As we now emphasize in the caption of Figure 1, the stimulus was indeed stopped at movement onset (defined above).

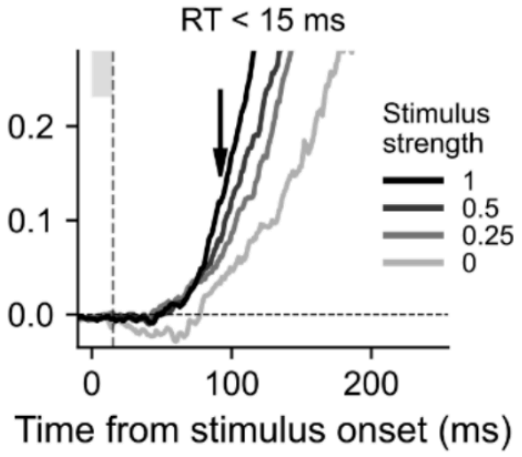
**Figure 1**. [...] The stimulus was stopped at the response movement onset.

Thus, the influence it has on the final choice in express responses is primarily attributed to delays in the sensorimotor processing pipeline. We also emphasized this in the main text:

This suggests that in a large fraction of trials, due to delays in the sensorimotor processing, rats update their response based on sensory information while moving from the central port to the selected lateral port.

3) perhaps indicate that the numbers in the legend (0 0.25 05 1) refer to stim strength if this is the case?

Thank you for the suggestion, we have now updated the legend.




4) Fig 2j: Some comment on why the prior is negatively correlated with position soon after stim onset and then positively correlated with position longer after stim onset would be helpful. One might expect the shift to be from high correlation with prior to less correlation with prior as stimulus information is integrated.

Our hypothesis is that the negative correlation before the movement onset arises because the rats are preparing to move to the side associated with the prior, placing their body slightly on the opposite side at fixation. We have now made this hypothesis more explicit in the main text.

The prior information, on the other hand, modulated the position of the animal even before the onset of the response trajectory (Figure 2j, right panel). Interestingly, such correlation was negative before the rat left the central port and became positive immediately after. This suggests that rats prepared their response while inside the central port, by placing their snouts opposite to the side they planned to choose.

5) p.12-13 Model explanation is confusing, e.g. what determines “x1” and what determines the time of “threshold collapse” which seems directly related to the motor initiation time? The theta_DV threshold is not clearly defined – is it the same as “theta” on the graph in Fig 1A? But then it appears that the level “theta” (or -theta) is never reached because the thresholds collapse. Perhaps that is because an “Action Initiation” process has reached threshold, in which case **for clarity it is essential that the process is included in the figure** (and I apologize if it already is!). If there are these two types of methods for initiating action (accumulation to theta_DV *or* action inititation) then I think both methods should be shown and explained clearly in the figure.

We thank the reviewer for this suggestion. We have now improved the explanation of the model and added a trace on top of Fig. 5a that represents the urgency process (that we termed Action initiation process in the methods), triggering the rat movement when reaching its corresponding bound.

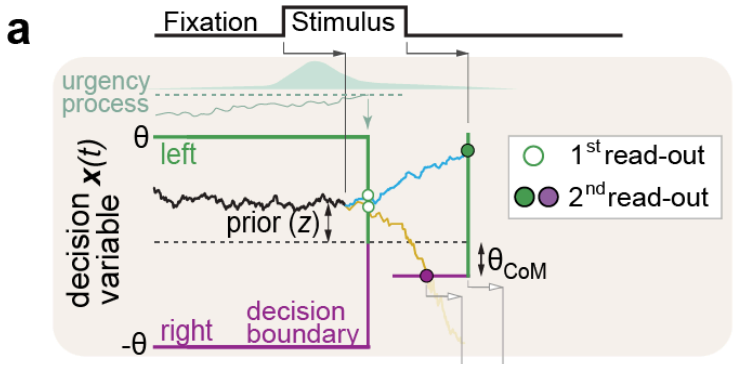



**Figure 5. [...]** the onset of the response movement is triggered by a stimulus-independent urgency process (turquoise)

We also modified the model explanation to make clearer that there are two different processes that can trigger the response:

Crucially, the decision-making module includes an urgency process (Figure 5a, top) and can initiate the movement independently of the decision variable, thus accounting for the prevalence of express responses (Hernández-Navarro et al. 2021; Hawkins and Heathcote 2021) (Figure 1c). When the urgency process triggers the response movement onset, the targeted port of the initial trajectory is set by the sign of the decision variable at this moment (x1). We refer to x1 as the first read-out (Figure 5a, empty circle). Importantly, the vigor of the initial trajectory depends on the value of x1: the more evidence towards the selected port, the faster is the trajectory (Figure 5b). Accordingly, when the response movement onset

is triggered by the decision variable hitting one of the decision bounds ($x1 = \pm\theta$), the initial response has the maximum vigor.

6) p.20: on the similarity of strategy suggesting commonality over **75 million years, this sentence ignores the (commonplace) occurrence of convergent evolution**.

We agree with the reviewer and have amended this paragraph to lower the claim, discussing the two possible explanations for the commonalities found between rats and humans:

The behavior of rats and humans was very similar, despite performing very different types of motor response (full-body orienting vs. arm movements). This resemblance suggests that the ability to adjust movement endpoints and vigor while processing asynchronous information is a fundamental skill in the sensorimotor repertoire of vertebrates. At this point we can only speculate whether this skill evolved before the two species diverged around 75 million years ago (Striedter and Glenn Northcutt 2019) or if it arose independently. To this respect, it is worth noting that the basal ganglia, which regulates motor vigor (Turner and Desmurget 2010), is a highly conserved area since its emergence in the first vertebrates 500 million years ago (Grillner and Robertson 2016). Beyond their similarities, humans differed from rats as they performed many more reversals.

7) p. 24 The explanation of Action Initiation includes a **"Trial Index" factor in the drift rate – does that mean there is a monotonic change over trials in likelihood of action initiation?** "Trial index" is neither defined nor mentioned till later (Supp Fig 1), so worth adding something about it and why it is a factor in the model when first mentioned.

We agree that this was not clearly explained and have now included the following sentence in the Methods section (p24):

This linear relationship on trial index accounted for the general slowing of reaction times within session, as we have shown previously (Hernández-Navarro et al. 2021).

8) p.24 "until the stimulus is fully integrated": this is confusing – what determines whether it is fully or only partially integrated? I would have thought the magnitude of x2 determines how much it has been integrated and this is inextricably tied in with whether it has reached a threshold or no, so the "or else" is confusing. Do you mean until "stimulus offset"? But in that case no clear offset time has been designated as far as I can tell.

This sentence refers to the stimulus information that has not been processed when the animal leaves the port because of sensorimotor delays. As we mentioned in point 2, we have now made it clearer that the stimulus stops at movement onset and that all the sensory processing after that moment is due to sensorimotor delays. Further, we now improved the explanation of the model, editing the text and adding references to the equations that define the MT.

The model sets apart from traditional decision-making models by fully describing the dynamics of the orienting response, and not just its end point. It consists of a decision-making module and a motor module that reads the accumulated evidence at two discrete time points. The decision-making module extends the classical Drift-Diffusion Model (Good 1979; Wald 2004; Gold and Shadlen 2007; Roger Ratcliff and McKoon 2008; Urai et al. 2019; Gupta et al. 2023) whereby a decision variable, $x(t)$, initialized at a value proportional to the prior evidence, integrates sensory evidence over time. When one of two decision bounds is reached, the associated side is selected. Crucially, the decision-making module includes an urgency process (Figure 5a, top) and can initiate the movement independently of the decision variable, thus accounting for the prevalence of express responses (Hernández-Navarro et al. 2021; Hawkins and Heathcote 2021) (Figure 1c). When the urgency process triggers the response movement onset, the targeted port of the initial trajectory is set by the sign of the decision variable at this moment ($x_1$). We refer to $x_1$ as the first read-out (Figure 5a, empty circle). Importantly, the vigor of the initial trajectory depends on the value of $x_1$: the more evidence towards the selected port, the faster is the trajectory (Figure 5b; Methods Eq. 1). Accordingly, when the response movement onset is triggered by the decision variable hitting one of the decision bounds ($x_1 = \pm\boldsymbol{\theta}$), the initial response has the maximum vigor. In either case, due to afferent and efferent delays, there is still sensory information in the processing pipeline at the movement onset. This information is incorporated into the movement in a second read-out $x_2$ (Figure 5a, filled circles), that takes place once the whole stimulus evidence is fully integrated, allowing for the possibility to update the response trajectory. Specifically, the trajectory is sped up or slowed down if the evidence towards the initial choice increases with respect to the first read-out (i.e. $|x_2| > |x_1|$) or decreases ($|x_2| < |x_1|$), respectively (Figure 5c; Methods Eq. 2). When the decision variable accumulates enough evidence against the initial choice, hitting a specific bound ($\boldsymbol{\theta}_{CoM}$), the decision is reversed: a new ballistic trajectory is drawn towards the opposite port, implementing a change of mind (Figure 5a, bottom, yellow trace). Note that, while in the model we can identify all CoMs, only a subset of those will be classified as trajectory reversals using the detection method based on trajectory deflections as in rats. Trajectories, generated using the principle of minimum jerk which ensures smoothness, had only two parameters: the targeted port and the MT (Methods). Therefore, the updating of a trajectory came down to drawing a new trajectory, that smoothly continued the initial one, but in which either the remaining MT was shortened or extended (vigor update), or the final target was switched (CoM). Although relatively complex, most of the properties of the model are derived from our previous observations (see Table 1 for an explicit link between observations and model features).

9) p.26 Within Alternative Models it is surprising that the authors did not assume, instead of “random initial choice” rather “initial choice equivalent to the prior” as that seems less of a “straw man” model for comparison. In fact from the diagram shown in Fig. 5 with such a short time between onset of stimulus integration and the collapse of the decision bound it seems that the initial choice would always be based on the prior. So a bit of explanation is needed as to how often that does and does not happen i.e. how well the data is well-fitted by a model that simply randomly commences a movement in the direction indicated by the prior.

We thank the reviewer for this suggestion. We covered this point in our response to point 1 above.

There are quite a few grammatical errors of English, particularly in the Introduction and Discussion, which could easily be corrected by sending the text through a GLM. e.g., in the first few pages:

p.2 “allows *us* … to” or one can say “allows for the prediction of …”
p.3 “Subjects … *their*”
p.3 “reverting *to* the”
p.6 “of a camera”: I assume taken “by a camera” and “of the animal”?
p.6 “consisting *of*”
p.6 “impact … on trajectories”

Thank you for these corrections, which have all been corrected.

## Reviewer #4 (Remarks to the Author):

Dear authors,

First, my sincere apologies for the late review. This is a topic close to my heart and as such I wanted to read it thoroughly but initially lacked time and focus.

This paper is a joy to read. You systematically combine research fields that have mainly been in parallel universes. I appreciate how you bring continuous (actually intermittent) updating of movement trajectories and especially vigor into the domain of decision making.

The experimental results seem crystal clear, although you do need a couple of data analysis tricks that require **better explanation**. I do understand how you introduced biases using P_rep of stimulus, but **how you converted this into 'z'** remains unclear to me (I did read methods and supplement but would need the earlier publication to really get it. Can you give a bit more of an intuition? - Pg 3 does not suffice with 'logistic regression'). So, I am very supportive to your study, but would like you to convince me a bit more with **statistical tests and proper identifiability of the model**. Also a **deeper discussion about the pro's and con's of the model (it is rather descriptive in place) would appreciated**.

We thank the reviewer for her/his kind comments and thoughtful suggestions, that we addressed point-by-point below.

1) Pg 3: The ref. to Hermoso-Mendizabal suggest that you might have re-used that dataset. I don't believe you did. **Make clear that this ref. is about an analysis trick**, not about this dataset.

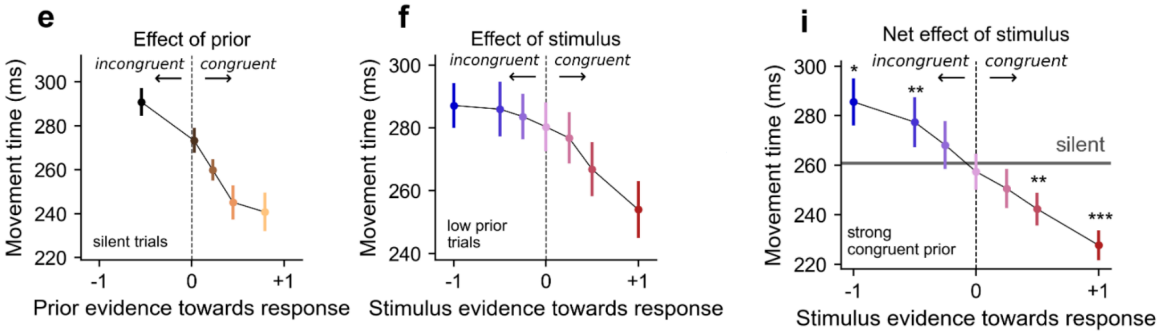
The experimental dataset used in this work is indeed different from that of Hermoso-Mendizabal et al. The reference is to indicate that the task is the same. We now cite the paper only when talking about the Repeating/Alternating structure of the trial

sequence. From the two groups of animals used in this study, Group 2 was already used for a different purpose in Molano-Mazón et al. 2023 and Duque and de la Rocha 2023. We now indicate this in the animal group table we have in the methods.

2) Figure 1F and 1I, **please indicate** in the panels the difference. Now you can only infer with hard work from the caption the difference between the two.

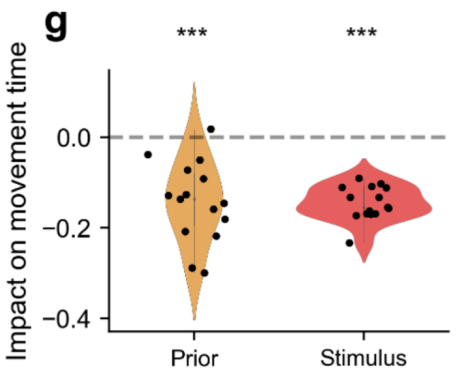
Thanks for this suggestion. We have now added titles and text to these panels so the figure can be understood more easily. We have also added the following text to the caption of Figure 1 (panel i)i:

Average movement time as a function of stimulus evidence supporting the decision, in express response trials (RT<50 ms) with prior evidence supporting the decision (prior evidence towards response z > median) - note that this is a different subset of trials from panel **f** which used only low prior trials.



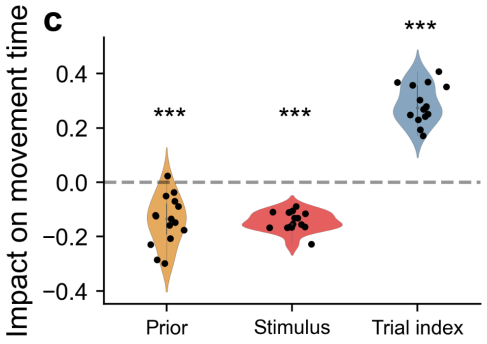

3) Figure 1G: **What is the quality (r2 for example) of these regressions?** The is clearly a significant influence, but how much variance remains unexplained (in the supplement you do indicate that there is a trial effect that you also regress out).

Thanks for the suggestion. We have now performed a t-test for each regressor weight across animals and report their p-value in fig. 1g (<0.001). We further computed the $R^2$ for each animal and now report the average $R^2$ in the caption of supp. fig. 1c. The regression model with prior, stimulus and trial index regressors explained 15 percent of the variance of the movement time of the rats on average.
Fig. 1g now looks like this with additions in caption:




c) Regression weights of the movement time against the prior and stimulus evidence towards the response (mean $R^2$=0.15). Each dot represents one animal. Asterisks correspond to a t-test against the null hypothesis that the mean of the distribution of weights is not different from zero. ***: p<0.001.

Supplementary Fig. 1c looks like this:




**c)** Regression weights of the movement time against prior evidence towards the response, stimulus evidence towards the response, and trial index (mean $R^2$=0.15). Points represent individual animals. Asterisks correspond to a t-test against the null hypothesis that the mean of the distribution of weights is not different from zero. ***: p<0.001.

4) Pg 2: **Visser et al. 2023 JNP** have also shown changes in the unfolding trajectory and looked into stretch reflexes. May go with references to Resulaj and Van den Berg - also in other places.

Thank you for the reference. We have added it at various places in the manuscript. We have also extended the explanation of CoMs/reversals to make clearer the difference between the two.

Previous work have shown that fluctuations in the stream of sensory evidence can lead to changes of mind (CoMs), i.e. an internal change in the decision variable towards the option that is aligned with the most recent evidence (Kiani et al. 2014; Peixoto et al. 2021; Boyd-Meredith et al. 2022;). In some studies in humans, CoMs could be identified by reversals of the response trajectory or muscle activity (Resulaj et al. 2009; van den Berg et al. 2016; Visser, Medendorp, and Selen 2023).

5) Pg 4: Nice description of finding with proper reference to the figures, but I miss any form of **statistical evaluation!** I know that we are moving away from P-values, but some form of statistical inference should at least be provided.

We agree that proper statistical tests were missing here. On top of the statistical test on regressors weights described above (point 3), we have now added p-values from t-tests comparing movement time in silent vs non-silent trials, separately for different stimulus evidence values (Figure 1i). The tests confirmed that indeed stimuli congruent with the response speed up the response trajectories while incongruent stimuli slow them down. We changed the description in the Results section to reflect this:

MT decreased for congruent stimuli, but increased when the stimulus was at odds with the final choice ($p< 0.01$ for both incongruent and congruent stimulus, Figure 1i)

Moreover, we have added the outcome of these statistical tests in both Fig. 1i and its caption:

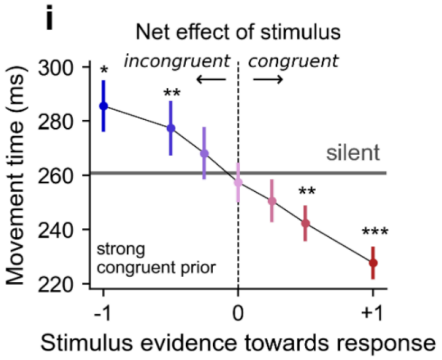



Figure. 1 [...] Asterisks denote paired t-test of MT silent vs non-silent trials across rats. *:$p<0.05$, **: $p<0.01$, ***: $p<0.001$.

6) Figure 2D: **how was movement onset defined?** Might be hidden far in the supplement, but would be nice to already mention.

Good point. We have now added a definition of the movement onset to the main text. We no longer call it “movement onset” but “response movement onset” because rats kind of move all the time and we wanted to be clear that this refers to the movement associated with the response trajectory.

In our experimental paradigm, the response movement onset (defined as the time at which the rat leaves the central port), was often driven by a proactive process, leading to a large fraction of reaction times (RTs)

7) Splitting time calculation on page 8 (panel 2H): give at least a hint, detail can stay where they are.

Thanks for the suggestion. We concur that the definition of the splitting time was a bit opaque. We have now included a short sentence in the main text that defines more clearly how the splitting time is computed.

To finely quantify the onset of this impact, we identified the earliest time at which the stimulus strength affected the trajectory. Precisely, the splitting time was defined as the first time point *t* starting from stimulus onset where the correlation between position *y*(*t*) and stimulus strength became significant (Figure 2h; see Methods).

We have also extended the explanation of the splitting time in the Methods:

Pearson correlation across trials between the position of the animal along the *y* axis and the stimulus strength. Defined as the changes caused by different stimulus strengths *within* the same stimulus category (Left or Right), the Splitting Time was not influenced by differences in the trajectory between Left and Right stimuli and thus it was not contaminated by the prior. In other words, we used the stimulus strength, and not the (signed) stimulus evidence, because it was completely independent (i.e. unpredictable) from trial to trial. We extracted the time-series of p-value for these correlations and identified the splitting time as the last non-significant time point from stimulus onset using a criterion of p=0.05.

8) Figure 3B and 3E: I guess you present in B the actual movements (including direction) and in E you have the chosen port always positive. Make it clearer. Why are trajectories much smoother in E than in B?

We thank the reviewer for this suggestion. Panels **b** and **e** correspond to individual trajectories (from a single session) and trajectories averaged across trials (for each animal), respectively. We have now added titles to the panels to make this more explicit. Moreover, panel **b** only includes trials that start with movement towards the left port as indicated in the caption ("Only 100 trials randomly chosen from the left-response no-reversals and right-response reversals in an example session are shown"). In **e** we flipped the rightward responses along the y axis before averaging. We now explain this in the caption:

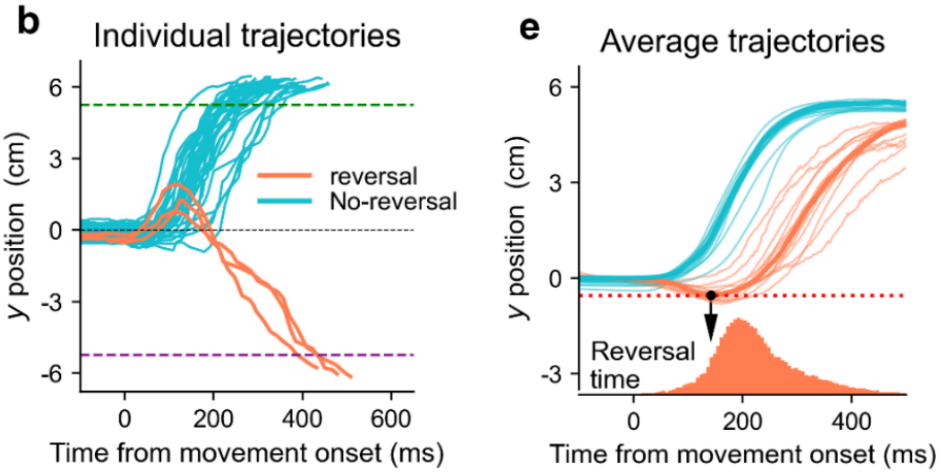
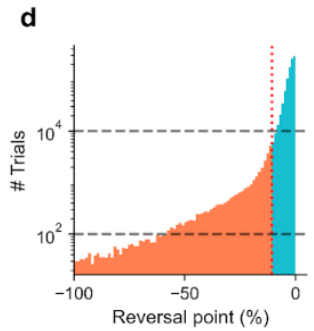


**Figure 3. [...] b)** Individual reversal trajectories (orange traces) and standard trajectories (i.e. no reversal, blue traces) from one single session, showing the position of the animal snout along the y-axis across time, from an example session. Response ports are represented with dashed lines. Only left-response no-reversals and right-response reversals in 100 randomly chosen trials from an example session are shown. **[...] e)** Average trajectories across trials for reversal and standard trajectories aligned at movement onset for individual animals (thin lines) and averaged across animals (thick lines). Trajectories corresponding to rightward responses were flipped along the y axis so that all trajectories ended at a positive value. Bottom: Distribution of the reversal time defined as the time from motion onset to the reversal point.

9) Figure 3D: **I don't understand this figure**. Explain better.

We apologize for the confusion. We have updated the caption with more details:




**d)** Histogram of the reversal points, defined as the maximum distance from the central port toward the non-selected port for each trajectory, and expressed as a percentage of the total distance between the central and lateral ports. Reversal points exceeding the detection threshold (vertical red dashed line) are categorized as trajectory reversals and highlighted in orange. The data includes all trials across all animals. Notice the logarithmic y-scale.

Page 10 "We explicitly inserted an **urgency component**" -> did you pilot and not find anything like the rats without? If so, indicate.

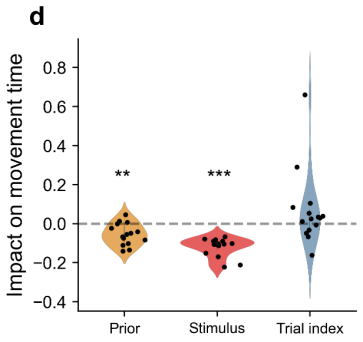
Although we never tested removing this constraint, we first piloted the experiment with a longer response window (500ms instead of 300ms) and higher stimulus evidence. We observed that the impact of the prior on the subjects choice was small, and for this reason we decided to reduce the response window and limit the amount of stimulus information provided to the subjects. Here is the text in the Methods in which we explain the rationale behind the paradigm for human subjects:

Subjects initiated each trial by maintaining the index finger of their preferred hand pressed against the starting point for 500 ms (fixation period). This adjustment of the fixation period was implemented because human subjects, unlike rats, were less accustomed to the task and required additional time to prepare for the trial. [...] Importantly, subjects had to start moving the finger within 300 ms after stimulus onset, otherwise the trial was considered a

missed trial and the screen turned yellow. This time constraint introduced an urgency component which nudged participants into initiating their response early based on prior expectations (matching in this aspect rats' spontaneously impatient behavior).

Again, for the human subjects you report figures, but no **statistics or comparisons**.

Fair point. We have now estimated the same regression model of the movement time as for rats (new supplementary Figure 3d), and report the average $R^2$ as well as p-values for regressors weights across participants. This showed that prior and stimulus evidence had a significant impact on MT, as in rats, whereas trial index did not (probably because it did not affect motivation the same way as it did in rats).




**Supplementary figure 3. [..] d)** Regression weights of the movement time against prior evidence towards the response, stimulus evidence towards the response, and trial index (mean $R^2$=0.07). Points represent individual subjects. Asterisks correspond to a t-test against the null hypothesis that the mean of the distribution of weights is not different from zero. **: p<0.01, ***: p<0.001.

First, I like the simplicity of the model, with an interaction between decision variable, urgency signals and movement trajectory. Only with the supplement in. hand was I really able to work out what Figure 5 describes, but I don't see how to make it simpler. Nevertheless, **I would like you to be even more clear about the fact that your model is intermittent in selecting and updating trajectories**.

Thank you for this suggestion. We have now modified the presentation of the model to emphasize that it has two read-outs:

It consists of a decision-making module and a motor module that reads the accumulated evidence at two discrete time points.

We have also added information about the urgency signal to Figure 5.

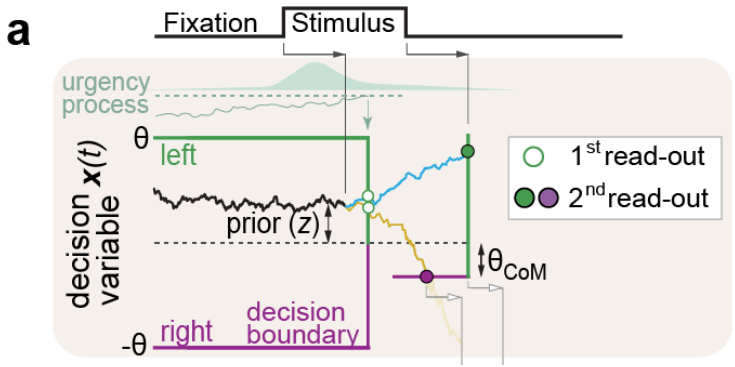


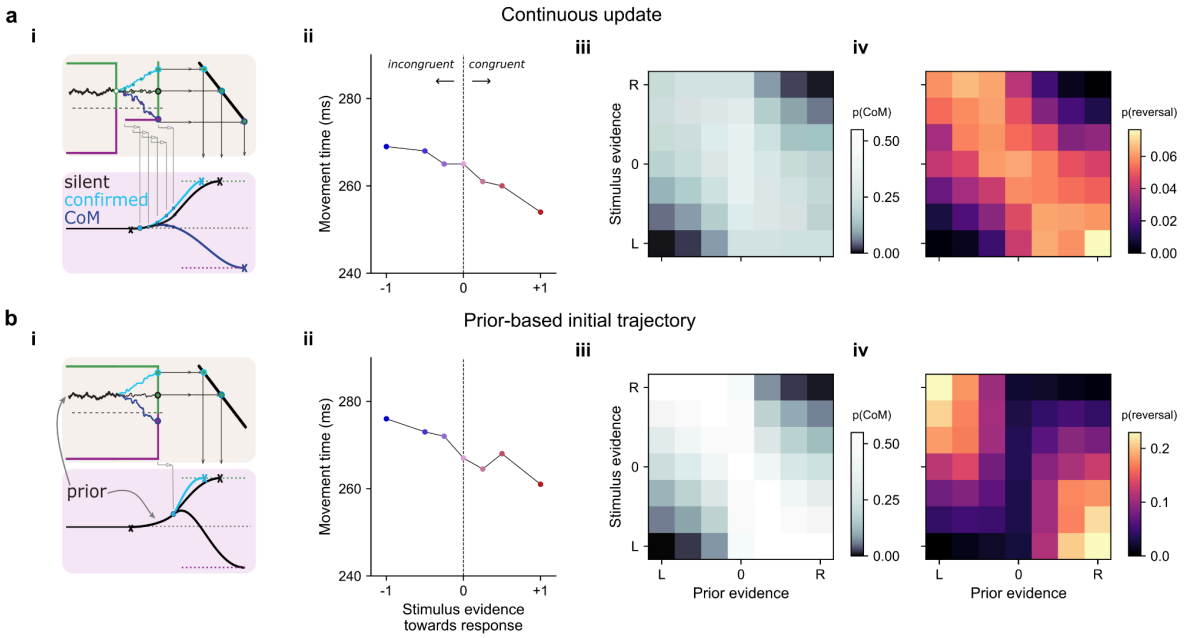

**Figure 5. [...]** the onset of the response movement is triggered by a stimulus-independent urgency process (turquoise)

Also, **the choice of MT** is not based on a process, it is assumed to be known (driven by 'vigor'?).

We thank the reviewer for this question. We realized that the way the MT is generated in the model was not completely clear in the manuscript. The MT is based on the first and second readouts as described in Fig. 5c and d, respectively. We have now changed the explanation in the Results where we make explicit reference to Eqs. 1 and 2 in the Methods. In these equations, it is shown that MT is a stochastic variable that depends on the accumulated evidence (via two readouts, x1 and x2) and in the animal “internal state” through the trial index. The text in Methods around these equations has been also edited to make it more clear.

I know that most OFC models also do not have an emerging movement time, but they are at least not intermittent. **Please discuss why you choose intermittent updating (for computational feasibility in part?).**

This is an important point. Whether updating of trajectories is intermittent or continuous has been an object of debate in the literature, where experimental data has often been shown to be consistent with one of the two possibilities, but to our knowledge no decisive evidence has settled the debate. Unfortunately, our study is no exception. We chose to have two updating times because it is the minimal number of updates you need to explain the data, and hence the most parsimonious (and computationally tractable) compared to those that have three, four, etc or continuous updating. Nevertheless, we have now included an alternative model implementing a continuous update of the decision variable for comparison, and we show that it accounts for the experimental data as well as the two-readout model (supp. fig. 12a).




**Supplementary figure 12.** Performance of alternative models. We evaluated the vigor and the trajectory reversals in two alternative models: a model where the vigor in the trajectory is continuously updated by reading out the evolution of the DV (instead of reading the DV in two discrete time points) (*Continuous update*, a) and a model where the initial trajectory only depends on the prior evidence instead of on the first readout of the decision variable (*Prior-based initial trajectory*, b). **i)** Cartoon illustrating each model. **ii)** Median movement time in express response trials (RT< 50 ms) with small prior (|z| < 10% percentile), as a function of stimulus evidence towards the selected port. **iii-iv)** Probability of CoM (iii) and reversals (iv) as a function of prior and stimulus evidence.

We believe that the intermittent-vs-continuous nature of trajectory updating can only be addressed by designing new experimental paradigms taylored for this purpose (i.e. where the decision is based on at least a handful of independent pieces of sensory evidence distributed in time during the trajectory). We now extended the discussion about this point:

One open question is whether more updates occur along the trajectory (Enachescu et al. 2021; Christopoulos, Bonaiuto, and Andersen 2015) or if the brain continuously transforms the decision variable into changes in the response trajectory (Friedman, Brown, and Finkbeiner 2013; Lepora and Pezzulo 2015). We found that a continuous-update model accounted for the experimental data equally well than our two-read-out model (Supplementary Fig. 12a). These two scenarios could be differentiated by assessing whether novel sensory evidence is integrated into the trajectories with a fixed delay (as predicted by the continuous hypothesis) or only at two different time points (as predicted by the two-readout hypothesis). However, although in our paradigm sensory evidence fluctuated within each stimulus, animals typically relied on only one or two stimulus samples (~100ms), and thus this analysis does not provide sufficient evidence towards either type of update scheme. A paradigm where stimulus evidence is resolved at a finer time scale, such as the click task (Brunton et al., 2013 Science), would be more adequate in that regard. Nevertheless, it is worth noting that the implementation of such a continuous model would require accommodating the dynamics of evidence accumulation, which typically exhibit fast

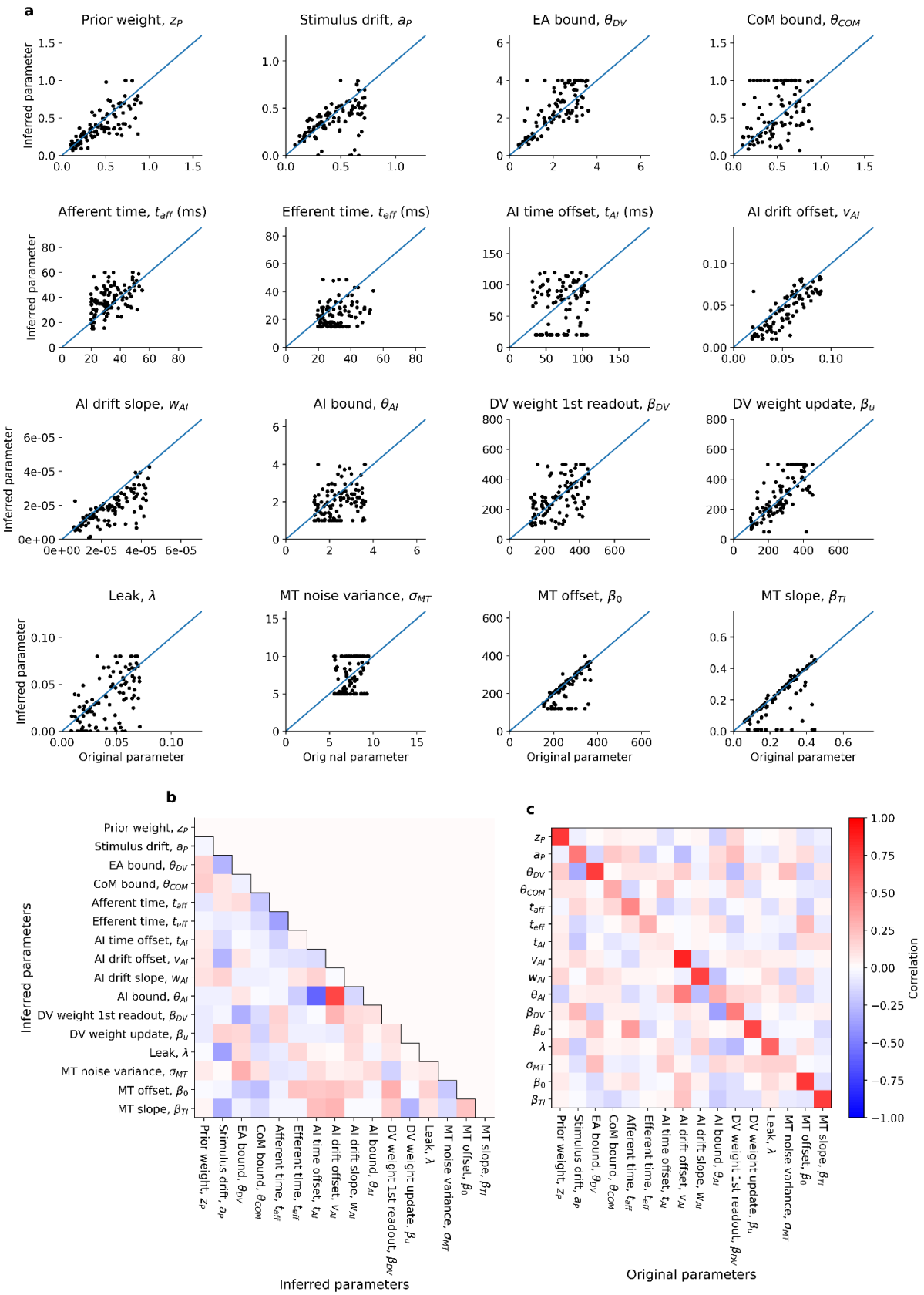
sudden changes in the decision variable, to the constraints of optimal motor control which imposes smooth, gradual changes in velocities and acceleration.

I like how you use sophisticated techniques to determine the parameters of the model. I am not sure this should go into Figure 5. It could be a methodological supplement.

We agree that including this in the main figure is debatable, but since using a simulation-based approach for such complex behavioral models (with more than a dozen parameters and stimulus conditions, as well as three output variables) is novel (to our knowledge), we would prefer to maintain it.

I appreciate how you find the parameters and construe a number of checks on the fitting procedure. However, the Approximate Likelihood in the supplement is not that convincing (Supplementary Figure 4, red vs blue). You might say that supp. Figure 5 provides proof, but that is only the stability of finding the same solution over-and-over again. I would have liked to see a **response dataset** based on a known generative parameter set and then fit and compare the posterior set.

We thank the reviewer for this suggestion. We performed a parameter recovery test where the model was first simulated for different ranges of parameters and then parameters were estimated for each synthetic dataset using the approximate likelihood. The estimated parameters are compared to true parameters and with each other. We now show the results in supplementary figure 16. We found that most parameters were well identified using the same estimation procedure as we used for experimental data. Some parameters were more tightly constrained by the data than others.




**Supplementary figure 16.** Assessing the model identifiability with a parameter recovery test. A simulated dataset was generated based on a known parameter configuration (original) and then the parameters were inferred using the fitting pipeline. This is done 100 times with random parameter

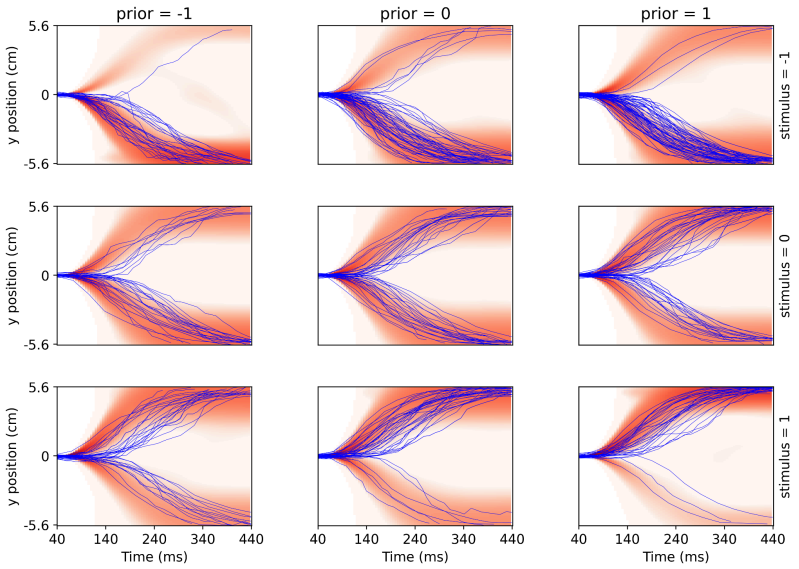
configurations. **a)** Inferred parameters vs the original. The blue line represents y=x, indicating perfect recovery. **b)** Correlation matrix between the inferred parameters. Since the matrix is symmetric, only the values below the diagonal. **c)** Correlation matrix between the inferred and original parameters. The diagonal corresponds to the correlations of the variables in **a**. Although the majority of the parameters show relatively good recovery, in some of them the recovered values differed or tended to fall on the bound of the permitted interval: $\theta_{COM}$, $t_{aff}$, $t_{eff}$, $t_{AI}$ and $\theta_{AI}$. The latencies $t_{aff}$, $t_{eff}$ show some degree of degeneracy, illustrating the difficulty in separating them based on the behavior of the animal. However, the recovery of the sum $t_{aff} + t_{eff}$ was precisely estimated (the correlation between the sum of the original and the inferred $t_{aff}$, $t_{eff}$ was 0.81). The bound $\theta_{AI}$ and $t_{AI}$ were partly degenerated. Finally, regarding $\theta_{COM}$, it is possible that the small proportion of CoMs and their indirect impact on the fitted variables (choice, RT, and MT) hindered the capacity of the fitting pipeline to fully recover this variable in some cases.

We have also added a couple of sentences to the methods explaining the recovery test and the results.

To further assess the identifiability, we performed a recovery test whereby a simulated dataset was generated based on a known parameter configuration (original) and then the parameters were inferred using the fitting pipeline. We found that most parameters were well identified, although some of them were more tightly constrained by the data than others (Supplementary Fig. 16).

Also, on the real data the model does not completely capture the kinematics (well that is also not expected given the simplicity of the simulated minimum jerk trajectories), especially in **Figure S6** all experimental traces are later and more shallow than those from the model. Please comment (I still like the model very much, but want you to reflect on the aspects it does not capture well).

This is an interesting point. The offset between experimental and simulated trajectories emerges partially because rats need to completely remove their snouts from the central port before starting moving towards one of the lateral ports (see Fig. 2b and c). Because the photogate in the port is *not* in the border but inside the port, this introduces a delay between the moment we record they come out and the moment they can freely move along the y-direction. We found that adding an offset of 40ms between the reaction time (i.e. initiation of movement along the x-direction) and the initiation of movement along the y-direction leads to a much better fit:

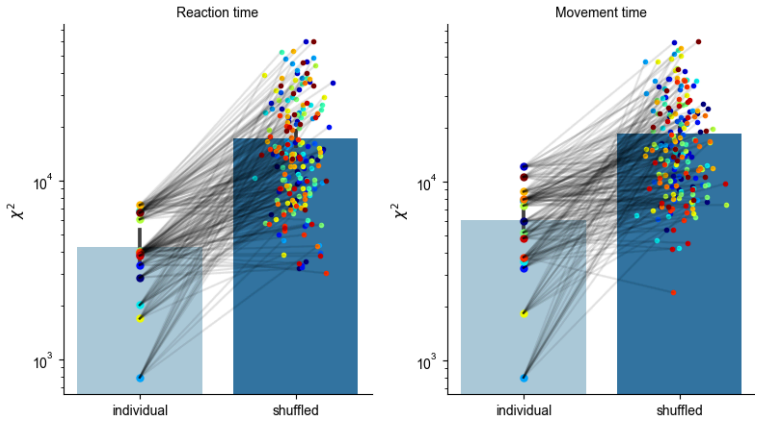



However we decided not to include it in our results because this offset does not affect the main claims of the study. We added a note to the caption of Supplementary Figure 4:

The lag of the experimental trajectories with respect to the model predictions is likely present because rats need to completely poke out from the central port before starting moving towards one of the lateral ports. Because our port photogate is inside the port and not in the outer border, there is necessarily a short delay between the initiation of movement along the x-direction (poking out) and the y-direction (moving towards one response port).

In the supplement you also show density distributions for MT, RT and proportion choices (the latter only for the model, **can you compare these to the actual ones**). I would have liked it if you would for example **shuffle the dataset** (link with simulated data from another animal) and show that you do much better than chance.

We thank the reviewer for this suggestion. We have now performed the following analysis to evaluate the capacity of our model to capture the individual variability in rats: for each rat we computed the $\chi^2$ between the experimental reaction time distribution and that produced by the model fitted to that same animal. We then computed the $\chi^2$ between the same model and any of the other 14 rats (shuffled data). We then repeated the same analysis for movement time instead. We copy the figure below:




For both the reaction and movement times, the $\chi^2$ is significantly lower than the values obtained by shuffling, in all 15 rats.
We have included the above explanation in the methods and added the following sentence in the caption of Supp. Figures 6 and 7 referring to this statistical test:

For each rat, the chi-square distance of the RTs/MTs between the data and the model was significantly lower ($p<0.05$) than the distance obtained after shuffling the animals' labels.

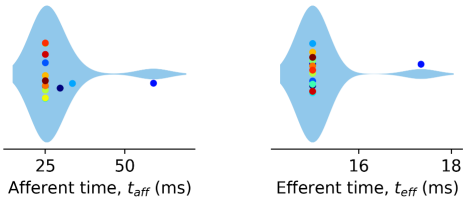
We are not sure we understand the reviewer's first point relative to proportion choices. We do show the proportion of reversals predicted by the model in Supplementary Figure 9, which mirrors the proportion in experimental data shown in Supplementary Figure 2.

Finally Figure S10 can be done much better. **Why not build (per rat) a corner plot of the parameter distribution**. This would show the consistency of the findings. The current figure is rather uninformative.

We agree with the reviewer that this figure was not very informative. Instead of a corner plot, we decided to show (see supp. fig. 16 above) a correlation matrix showing the correlation between the true and estimated parameters in a parameter-recovery test, as we believe this provides the relevant information in a more concise fashion.

Also, from the efferent and **afferent estimates it seems that they are discretely sampled**. Is that true, if so mention. If not, why are **so many point at the same value**. Can you reflect why sigma_MT reached the upper bound for all animals? What does that tell you about the model?

Afferent and efferent times tend to fall in the lower bound we imposed as a biological constraint during the parameter estimation procedure. As we discussed above (see reply to reviewer 2), this might be caused by the model failing to capture all the nuances of the behavior of the rats.




We have added a comment on the caption of Supplementary Figure 16:

The latencies $t_{aff}$, $t_{eff}$ show some degree of degeneracy, illustrating the difficulty in separating them based on the behavior of the animal. However, the recovery of the sum $t_{aff} + t_{eff}$ was precisely estimated (the correlation between the sum of the original and the inferred $t_{aff}$, $t_{eff}$ was 0.81).

Pg 21: **Only 6 rats had silent trials. Indicate in main paper**. These trials play a key role in your analysis, but for the majority of the rats you don't have them.

We now make this information more explicit:

To isolate the contribution of the prior evidence on vigor, we introduced, in a group of animals, *silent catch trials* where no acoustic stimulus was played (Group 2, n = 6 rats, mean±SD 6.7 ± 0.6% of all trials).

Pg 20: just before Making decisions in a continuous world: 'indicating a lack of bias', can you clarify?

We have now clarified this sentence:

Note however that the boundary for changes of mind was close to null in most animals (Supplementary Fig. 13). This indicates that rats did not require a larger amount of evidence towards the opposite response to change their mind, suggesting a lack of confirmation bias towards the initial choice.

Pg 20: **Making decisions in a continuous world**: Not sure this is the paragraph, but I would like a discussion about intermittent and continuous control adjustments. For example. there is a 2021 paper by Enachescu, Schrater, Schaal and Vassilios that combine a neurodynamic field model with OFC to model trajectory averaging when making a choice. I believe this should be discussed in relation to your findings.

We thank the reviewer for the suggested reference, which we now cite and comment in the extended discussion about intermittent VS continuous read-outs (see also above).

One open question is whether more updates occur along the trajectory (Enachescu et al. 2021; Christopoulos, Bonaiuto, and Andersen 2015) or if the brain continuously transforms

the decision variable into changes in the response trajectory (Friedman, Brown, and Finkbeiner 2013; Lepora and Pezzulo 2015).

Pg 21: **Model of prior evidence**:
Sorry, I cannot decipher this. Probably if I would know or study the Hermoso-Mendizabal paper I would. Can you write this more clearly and maybe even provide a cartoon?

We have now clarified in the methods how the weights obtained when fitting the GLM to the behavior of the subject are converted into the time-resolved prior z:

After fitting the GLM model to the choices of each individual animal, the prior evidence *z* was defined in each trial as the weighted sum of the trial-history regressors (i.e. those corresponding to the previous repetitions/alternations and those corresponding to previous rewarded and non-rewarded responses occurring in the last 10 trials), weighted by their corresponding regression weights. This represents a subject-specific characterization of the history-based prior evidence on a trial by trial basis.

Pg 26, just before Alternative Models: '**avoid identifiability issues**'. Sure, but the you should also show that your model parameters are identifiable in the current version of the model (see comments above).

We have now run a recovery test. See above.

Same paragraph, **please explain all variables**. I can deduce most, but am quite well versed in this domain. For a PhD-candidate this is hard, if not impossible, to decipher.

Thank you for spotting some missing information here. We now explain each variable of the model in the methods.

In summary, the model contains 18 parameters overall: 5 parameters related for the Action Initiation process ($t_{AI}$, $v_{AI}$, $w_{AI}$, $\sigma_{AI}$, $\theta_{AI}$), corresponding to the time at which the Action Initiation process starts, intercept and slope that determine the drift as a function of the trial index, the sigma of the Gaussian noise added to the process and the bound of the process; 8 parameters for the Evidence Accumulation process ($t_{aff}$, $t_{eff}$, $z_P$, $a_p$, $\lambda$, $\sigma_{DV}$, $\theta_{DV}$, $\theta_{COM}$) corresponding to the afferent and efferent delay, the weight of the prior, the drift, the leak, the sigma of the Gaussian noise added to the accumulated evidence at each timestep, the bound of the process and the bound that the evidence has to reach to change its initial choice, respectively; and 5 parameters for the response trajectories ($\beta_o$, $\sigma_{MT}$, $\beta_{DV}$, $\beta_{TI}$, $\beta_U$) corresponding to the intercept, the sigma of the Gumbel distribution, the weight of the evidence and the weight of the trial index, respectively, in the function that determines the

initial movement time in Equation 1, and the weight of the evidence difference in the function that determines the update in the movement time in Equation 2.

We thank the reviewers for their kind comments and address the few suggestions they have below.

Reviewer #1 (Remarks to the Author):
The authors have effectively addressed all my previous comments and those of the other reviewers through insightful additional analyses. The newly incorporated information is particularly valuable, enhancing both the clarity of the manuscript and the impact of the presented results. I recommend this manuscript for publication in its current form.

We thank the reviewer for her/his kind comments.

Reviewer #2 (Remarks to the Author):
I thank the authors for their clarifications and updates to the work. I am a bit disappointed that they were not able to share the code to allow us to review this alongside the manuscript, but leave it to the discretion of the editor whether this is mandated at submission.

As mentioned in our previous response, we completely agree with Reviewer 2 that providing others with the data and clear instructions for analyzing it and fitting the model is fundamental for our results to fully benefit the community. We have now created a source data file to reproduce all the main figures, two notebooks to generate the figures in the paper, and a .pkl file with all the raw data.

Minor: Perhaps indicate that you use 'RT' and 'MT' in a way that may be different with respect to other literature involving continuous rodent choices. For instance, https://doi.org/10.1101/2023.07.04.547681 uses the terms 'movement onset' to denote what you here call 'RT', namely the first movement made that does not yet indicate the animal's choice.

We thank this suggestion made by reviewer 2 because it made us realize that our definition of movement onset and reaction was a bit confusing. We now updated the paragraph to make it more clear:

In our experimental paradigm, movement onset, defined as the time when the rat leaves the central port, was often driven by a proactive process, leading to many trials where the stimulus duration was too short to trigger a response (Hernández-Navarro et al., 2021). Indeed, in 34.3 ± 12.2% of all trials (mean ± SD), the rat initiated a movement in less than 50 ms after stimulus onset (i.e. reaction time, RT < 50 ms). We call these short-RT responses express responses (Figure 1c).

Reviewer #3 (Remarks to the Author):
In the revised manuscript, the authors have done an exceptionally thorough job of addressing the prior critiques, for which I applaud them. I do think that between the editorial office and the authors a final grammatical sweep is needed (ensuring prepositions are correct in several places) but otherwise I have no concerns and think this is a very nice article.

We thank the reviewer for her/his kind comments. We have checked the grammar of the text and corrected the few remaining typos.

Reviewer #4 (Remarks to the Author):
I applaud the authors with this great study. All four reviewers had constructive and often similar comments, that were taken up in a constructive manner by the authors. The reply and paper are a joy to read. I am impressed how the authors build a continuous version of their model but clearly indicate that it is hard to disprove one or the other. I believe this paper is a great addition to the field and deserves publication.

We thank the reviewer for her/his kind comments. We hope making explicit the limitations of our results will inspire future experimental and computational research.