nature portfolio

Peer Review File

Alpha Frequency Shapes Perceptual Sensitivity by Modulating Optimal Phase Likelihood

Corresponding Author: Professor Vincenzo Romei

This manuscript has been previously reviewed at another journal. This document only contains information relating to versions considered at Nature Communications.

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

A version of this paper was originally rejected for publication by Nature Communications, however that decision was reconsidered after appeal by the authors.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
The authors conclude that alpha frequency influences sensory precision by modulating the distribution over phases.
There are a couple of conceptual and metholodological issues:
1. Can the hypothesized mechanism explain variation in observers' behaviour?
The authors suggest that higher alpha frequency increases the probability that the optimal phase will be present during stimulus presentation. Let's think this through: stimulus duration = 60ms. The authors observed an alpha frequency of 11.15 Hz (correct trials) vs. 11.12 Hz (wrong trials) i.e. a difference in alpha frequency of 0.03 Hz. Hence, the length of the alpha cycle decreases from 89.928 ms for wrong trials to 89.686 ms for correct trials. The coverage of alpha phases for wrong trials about 66.7% and for correct trials about 66.9% , i.e. a tiny increase in the coverage of alpha phases of 0.2%.
Can such a tiny increase in the coverage of alpha phases really explain a measureable behavioural effect? Can it explain a behavioural effect that can be measured reliably with only 150 trials per participant?
2. Is it possible to measure a change in alpha frequency of 0.03 Hz reliably given the constraints imposed by the experiment?
To estimate a 0.03 Hz difference reliably and obtain such an extremely high spectral resolution, one requires very long data segments. However, the segments are limited to only brief 700 ms prestimulus baseline.
Moreover, the estimation of the instantaneous phase with the Hilbert transform depends on the filter parameters. No data are shown that confirm the appropriateness of a filter range of IAF +/- 2 Hz.
The authors take the first temporal derivative to remove an aperiodic component. However, this is often not sufficient to remove 1/f noise as an important confound.
In addition to these two fundamental problems, there are a couple of other issues:
1. Changes in power which are a key confound are not reported.
2. No eye tracking data are reported. Eye movements are crucial confounds that need to be excluded.
3. The authors report only results at the sensor level. This makes it impossible to draw decisive conclusions about the underlying neural mechanisms. It is very likely that alpha frequency as measured on the scalp arises from multiple neural generators e.g. parietal and occipital sources which may have different frequencies. Hence, putative changes in alpha frequency may reflect changes in power of different alpha generators (rather than a change in alpha frequency).
4. The experimental paradigm involves crowding, detection etc. – possible effects of alpha frequency may act through various mechanisms making strong conclusions difficult.
5. It seems that there are only 150 trials per participant. How many trials fell into different categories (e.g. correct vs. wrong)? What is the impact of the correction procedure applied to the d-prime (for small number of trials, biases can be substantial)
6. Quite a few methodological details are missing e.g. visual angle, luminance, stimulus duration = 60 ms with a refresh rate of 85Hz - shouldn't that be a multiple integer? Parameters for Bayesian t-test? Statistical analysis of accuracy? Generalized linear mixed effects models are recommended. Fitting procedures and parameter recovery of HDDM?

Reviewer #2

(Remarks to the Author)
This paper extends current proposal regarding perceptual sampling by the human alpha rhythm by proposing that the relevance of alpha phase can be captured only at slower frequencies. This is based on the idea that with faster rhythms, it is more likely that a stimulus will fall under a 'good' phase of the alpha rhythm. The mechanism is elegant and consistent with previous proposals. The methods are sound and overall, the paper is well written, albeit with some typos. I have the following comments.

Abstract and line 388: "Bayesian statistics, and advanced computational models" – I think that the authors are referring to using the drift diffusion model, and the word "advanced" is not necessary or useful (DDM was described by Ratfcliff and colleagues in 1978). Why not just write "drift diffusion model" which is the same number of words?

Abstract, I would tone down the "unprecedented" from the "we provide… insights into the neural mechanisms through which IAF influences perceptual decisions" – I think that the seeds of the idea have been in some of the writings by VanRullen, Busch and others, as well as in some papers from the 1930's onwards, and we expect most things published in this journal to be "unprecedented", so it is fine without the hype words.

Abstract:
this sentence seems to be missing a word: "Higher probability for covering optimal alpha phases during same stimulus presentation, crucial for accurate perception, is intrinsic to faster than slower IAF." Perhaps "is intrinsic to faster rather than slower IAF"….

Methods:

Concerning the methods. It is established that alpha power, more consistently than frequency or phase, has an effect on visual detection. Additionally, peak alpha frequency often covaries with alpha power. Given the small magnitude of the effect reported it is imperative that the authors perform the same analysis they conducted by binning alpha frequency but by binning alpha power. In other words, the author should clearly show that their results are not driven by an effect in alpha power.
The manuscript seems to be a reanalysis of a previous dataset used in another paper published by the same authors (https://doi.org/10.1162/jocn_a_02026). If this is the case, the authors should clearly acknowledge this through the manuscript and summarize the previous analyses and findings.
Discussion:

Results
Criterion – perhaps there is some weak effect at specific time points that is missed by correcting over the whole time period? Other studies have reported criterion changes with alpha power, so it would be good to be careful about how this is worded.

Figure 4C – It is interesting to note that the significant bins for ITPC in slow IAF participants, which are not found for the fast IAF individuals, is in the high alpha to low beta range (appears to be 10-18 Hz approximately from the figure, at least). It would be good to point this out, rather than simply describing it as "confined to the alpha band" (line 294), which does not seem quite accurate.

I agree that drawing a clear connection between studies showing a role of alpha in detection (cited in the paper) and studies showing a role of alpha in temporal integration is a valuable theoretical contribution. In addition, the ideas put forward could be tested, and I think that it would be useful to make these implications clear with a few examples. For example, the role of alpha phase and IAF on detection should vary depending on stimulus duration (see Figure 5). Second, increasing IAF externally via entrainment/neurostimulation could improve detection while also decreasing the role of phase. Given that IAF varies naturally between people and as a function of age and neural disorders (such as SCZ as mentioned by the authors),

It would be good to be more precise with the use of the word "precision". For example: "In alignment with these findings, our results strongly contribute to this body of research as they indicated a significant role of IAF in dictating the precision of sensory acquisition, as spontaneous inter-trials fluctuations in IAF were able to strongly predict the accuracy of perceptual decision." As the authors note, they only measured accuracy and not precision of sensory representations since it was a simple detection (present/absent) task. It may indeed be the case that IAF might influence the amount of time that a brief stimulus is sampled during optimal phase, but the authors did not show that this is the case or that this decreased time necessarily reduces the precision of the neural response or perceptual interpretation. Again, this could be tested with a different paradigm, leading to a strong hypothesis that a task in which duration of sensory sampling improves precision is influenced by IAF and alpha phase.

The authors suggest that their model could explain null findings based on the fact that previous research has overlooked the importance of peak alpha frequency when looking at the effect of phase. Similar proposals have been advanced. For example, it has been shown that other intervening variables could explain the effect (or lack thereof) of alpha phase/frequency. Particularly, consideration of alpha power (Fakche et al. 2022 eNeuro) and/or aperiodic properties of the EEG (Deodato & Melcher, 2024, biorkiv) has been shown to affect result. These alternative viewpoints could enrich the discussion.

General: please use spellcheck for minor errors such as “temporal integratrion” and so on. I think that simply using Word to check the text would help.

Other comments:
Lines 56-57 if the authors were the first to postulate such hypothesis, they should add an appropriate reference. Otherwise appropriate references could be:
Valera et al. 1981
Kristofferson 1967
Lines 62-69 present an oversimplified view of the effects of neurostimulation on temporal integration windows. Many researches using entrainment or neurostimulation reported effects that are not entirely supporting this view. This should be acknowledged.
Lines 83-87 “This leads to the intriguing hypothesis that variations in alpha frequency play a more basic role in visual processing beyond its temporal aspects, influencing visual sensitivity and extending even to low-level stimulus features.”

It’s not clear how one thing would lead to another, the authors need to make a better point at this. This seems to be one of the key points in the paper. However, it is presented in an approximate way.

Lines 97-99. “This proposition, in turn, leads to another hypothesis closely intertwined: the ability of phase angles in dictating the accuracy of the response is expected to be more pronounced in participants or trials in which IAF is slower.”
This is true only for short lived stimuli. Markedly shorter than the length of an alpha cycle but not too short. The authors should specify in introduction and discussion the importance of stimulus length on the proposed mechanism.

Line 145 “perceptual decision-making perception” it’s likely a typo.

Lines 173-174. There seems to be an inconsistency between what the authors reported (-800 to -450) and the figure 1.

Lines 193-194, The effect is very small (it depends on the decimal places), this is concerning.

Lines 203-204. There is a statistic report but no indication over which channel/time-point or cluster.

Line 337. Integratrion is a typo

Lines 358-368. Again, the mechanism proposed is sound and consistent with previous evidences. However, the author should point that within the mechanism they propose these results should be expected only for stimuli with a specific length.
Figure 3 is missing the letter identifiers (A,B, and C).

Line 446: “and further ampliated” (while ampliated is indeed a word historically it is rare to find it in regular discourse)

Reviewer #3

(Remarks to the Author)
This study presents a novel and interesting view on the mechanisms underlying alpha oscillations’ effect on visual perception. They use a visual detection task and investigate how instantaneous and individual alpha frequency and alpha phase impact perceptual performance. The study demonstrates that individual alpha frequency (IAF) significantly influences perceptual accuracy, with faster IAF predicting higher sensitivity and accuracy in sensory acquisition. This relationship is moderated by the phase of alpha oscillations, where individuals with slower IAF exhibit more pronounced phase-related variations in perceptual accuracy, whereas those with faster IAF cover a broader range of phase angles, reducing phase influence on decision accuracy. These findings provide critical insights into previously reported conflicting results regarding alpha phase and perception. The authors used state-of-the-art analysis methods and computational modelling of the behavioral performance. They clearly present the current state of the field and the shortcomings of current theoretical models of alpha oscillations in perception.
I consider that this study is of very high quality and could have an important impact in the field. I have some concerns with the clarity, level of details reported and some of the analytical choices and controls. I list my comments below, roughly in order of importance. I believe that most of my remarks and questions will not change the main conclusions of the paper but hopefully improve its impact and quality.

1. The DDM analysis is interesting but some aspects should be explained in more details and better motivated:
a. The task's temporal constraints (none in this study) should be indicated before presenting the DDM results as such constraints heavily impacts DDM fitting, e.g. a relatively strict response deadline will impair DDM fit as the RT distribution is truncated. The authors explain how RTs were selected for this analysis in the methods (and their choices seem valid), but providing more information on the task structure (at the beginning of the result section or in the DDM section) would help readers interpret the findings.
b. The authors mention that it takes into account RTs but without elaborating on how RTs are integrated in their proposed mechanistic view of how IAF and phase shape perception. See also my other comment below on the lack of information on RTs in this study.
c. The authors should elaborate on what the DDM analysis adds on top of the SDT analysis as the two indices reported from the DDM (drift and bias) are conceptually very close to the SDT ones (sensitivity and criterion). Or whether the DDM analysis was used as a conceptual replication of the SDT findings.
d. The DDM commonly allows to estimate two other parameters not reported in the current study: bound and non-decision time. These parameters are not reported in the current study. Moreover, in the original paper presenting the DDM model used in the current study, Wiecki et al. (2013) explain that one should to allow all DDM parameters to vary according to conditions (here IAF) and perform model comparisons to determine which parameters should be fixed and which ones should be allowed to vary with conditions. I do not believe that this procedure was followed in the current study.
e. The authors do not describe how the diagram representing the DDM model (Fig. 2C) was obtained. Is it a generic graphical representation of the DDM or were the distributions and drift arrows simulated based on the model fit to the data? Also, the diagram should be described in more details in the caption.

2. The behavioral task is not described in sufficient details:
a. The stimulus composition is unclear. If I understand correctly, every stimulus was a checkerboard composed of black and white cells in catch trials, and with superimposed grey circles in target trials. Was a target composed of the checkerboard with a circle in every cell or only a subset of cells? It seems like it was in every cell but this is not explicitly mentioned in the methods.
b. What was the size of the entire stimulus (in degrees of visual angle)? What was the size of each square/what was the spatial frequency?
c. In the only depiction of the stimuli in Fig. 5, it seems like circles placed on white versus black cells had different shades of grey, what was the difference between these circles' colors? What specific feature of the circles was titrated in the staircase?
d. The staircase procedure is not explained, please provide more information on the type of procedure used (1-up-2-down? Quest method?), how many trials were included in this staircase, etc.
e. No information is provided on reaction times in this task (neither descriptive statistics nor any statistical test) although this performance index is used in the DDM analysis. This information is essential to evaluate the paradigm in general and some of the findings, and should thus be presented at the beginning of the results section (at least descriptive stats and/or graphical representation).

3. In general, I think the task should be explained in more details at the beginning of the results section and accompanied with graphical depiction of the trials time course, and together with descriptive statistics (and/or plots) of accuracy and RTs.

4. Because (one of) the main axe(s) of the study is individual differences in IAF, the authors should document this further.

a. It would be informative to know how much IAF varies within an individual and whether this variability is correlated with the individual's average IAF (e.g. is prestimulus IAF more/less variable in individuals with higher average IAF?). One inspiration for representing this variability could be Fig. 2 of Grabot & van Wassenhove (2017) (https://journals.sagepub.com/doi/full/10.1177/0956797616689369?journalCode=pssa) investigating interindividual differences in temporal-order-judgments. Of course, this is just a suggestion for graphically representing this variability.
b. Moreover, since objective target discriminability was titrated for each participant, and according to the authors' claims, there should be a correlation between participants' average IAF and their titrated contrast reached during the staircase. Showing this correlation would strengthen the authors claim on the impact of IAF on perceptual performance.

5. Some details are missing in the EEG methods:
a. I guess the authors only inspected occipito-parietal electrodes on the midline and right side because the stimuli were always displayed on the left visual hemifield. This should be explicitly stated in the methods section.
b. The authors omit important details on the EEG recording, namely: how many electrodes were used and what system/brand/device was used?
c. In the ICA preprocessing step: how many components were the data decomposed in? How many components were discarded per participant on average? The author mention that "components containing artifacts that could be clearly distinguished from brain-driven EEG signals", does that mean that all "non-brain-driven" component were discarded? In a typical 64 electrodes EEG dataset ICA decomposition, if the ICA decomposes the signal into 64 components, most components are not brain-driven, therefore it would mean that in this study a large number of components were discarded. Or did the author discard a specific and restricted type of components (e.g. blinks, lateral eye-movements, as is commonly done)?
d. The filtering should be described in more details (type of filter, cut-off frequencies, etc.) as this could potentially have an impact on oscillatory measures and would be necessary to replicate this study.

6. For the analyses described in the "Instantaneous alpha phase shapes perceptual accuracy" section, the authors use 2

phase bins and argue that this number of bins was chosen to have sufficient number of trials per bin. I understand aiming to retain enough trials per bin but I don't understand how the two bins were chosen. I am not an expert on phase analyses but since absolute phase of neural oscillations measured with EEG is meaningless (peaks and throughs depend on cortical folding and local or distant orientation of the dipoles contributing to the signal), wouldn't it be more sensitive/statistically powerful to choose phase bins individually using a data driven method? For instance, computing a phase opposition index between correct and incorrect trials? I might be missing a critical aspect of the analysis that would not be possible with another method, but I would like the author to better motivate this analytical choice.

7. The authors used 3 phase bins in the IAF analysis. They mention using the same procedure as another study but they should provide some evidence that the results do not depend on this specific setting. Do the result hold (at least qualitatively) when using 2 or 4 frequency bins? Moreover, the authors only compare the 1st and 3rd bin, I think this is a valid analytical choice, but the effect in the 2nd bin should be presented at least qualitatively or graphically. It would be valuable to know whether the effects observed in the 2nd bin are in-between the 1st and 3rd, or whether there is a non-linear relationship between IAF and perceptual performance.

8. The authors do not seem to control for alpha power in any of the analyses. Although the effect of alpha power on perceptual processes is still debated in the literature, this could be an important confound in the reported analyses, e.g. alpha power in trials grouped in the 1st and 3rd bins according to IAF values could be different and thus put the authors' claims into question. This should also be investigated at the interindividual level: is alpha power different in low versus high IAF participants?

9. The authors should at least mention which sensors were used in the analyses at the beginning of the results section. I understand that the details of sensor selection belong to the methods but knowing which electrodes were used for the analyses is fundamental to interpret the results. Moreover, in the results section they describe how electrodes were selected but it would be informative to provide the final distribution of selected electrodes across participants (i.e. how many times was each electrode selected). Also, if I understand correctly, only one electrode was used for each participant? This could be more clearly stated.

10. The authors do not discuss potential origins of individual differences in IAF. Although there might not be much to elaborate on in the literature (one potential factor I know of is white matter structure (Minami et al., 2020)), this is a central aspect of their study and I believe it would be interesting to at least propose future avenues for research on this question.
Minami, S., Oishi, H., Takemura, H. & Amano, K. Inter-individual differences in occipital alpha oscillations correlate with white matter tissue properties of the optic radiation. eNeuro 7, 2 (2020).

11. The 3rd IAF bin in Fig1 A & B seems to fluctuate rhythmically and the two significant clusters in Fig. 1A are roughly separated by a theta cycle (300-250ms). Could the authors comment on this? More generally, the authors do not elaborate on potential cross-frequency coupling although numerous studies have linked theta-alpha coupling in visual perception. I do not believe that this is mandatory but it would be interesting to discuss these potential interactions.

12. It would be interesting to discuss how the present findings relate to computational models of (alpha) oscillations (e.g. Karvat & Landau, 2024, although this specific study is concerned with temporal integration).
Karvat, G., & Landau, A. N. (2024). A role for bottom–up alpha oscillations in temporal integration. Journal of Cognitive Neuroscience, 36(4), 632-639.

13. In the "Instantaneous alpha phase shapes perceptual accuracy" section, no statistic or p-value is given. The authors should provide some information on the statistical results there.

14. There are a number of typos and errors across the manuscript (e.g. L.249 "invstgate", the citation of the DDM package used is Wiecki, not Wiekli, and it is not included in the references, etc.)

15. A perceptually homogenous color-scale should be used in Fig.4C (https://www.mathworks.com/matlabcentral/fileexchange/51986-perceptually-uniform-colormaps)

16. The title of section "Alpha-Phase clustering increases following correct responses only in slow IAF individuals." is confusing as the effect that is shown is post-stimulus presentation but not necessarily post-response (or at least we cannot know since RTs are not reported).

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
The authors have made significant efforts to revise their manuscript. However, several methodological concerns raised previously remain valid and require further attention. Below, I briefly outline the key issues along with recommendations for improvement.

1. The main conclusions still rely on computing a so-called ‘instantaneous alpha frequency’ by applying the Hilbert transform to obtain phase and then computing the derivative. However, this approach is flawed. The term ‘instantaneous frequency’ is misleading in the sense that frequency can accurately be computed only over finite time windows. Instantaneous frequency estimates are not consistent with fundamental principles in physics, as greater frequency resolution requires longer time windows. Michael Cohen who initially proposed this method showed in subsequently published research that this approach leads to biased estimates that depend strongly on 1/f noise and SNR. While the authors attempted to remove 1/f noise, their approach appears to be problematic. As shown on p.12 in their rebuttal, the PSD after 1/f noise removal shows substantially lower alpha power (an order of magnitude lower), which is rather surprising given the original PSD. To ensure unbiased frequency estimates, it is crucial to adhere to established methods such as FFT over longer time windows. This approach is widely validated and avoids the pitfalls associated with ‘instantaneous frequency’ calculations.

2. The authors suggest that despite the minute differences in IAF within each participant they observe substantial variability when computing variance pooled over time, trials and participants (p.2 in rebuttal). However, this approach conflates within trial variability across time, within-participant variability across trials and between-subject variability. Because the statistical inference (i.e. comparison: correct vs incorrect) is across trials within each participant, this conflation of variability is not appropriate.

3. In support of their own findings, the authors cite previous studies in high impact journals that reported comparable differences in alpha frequency. However, these older studies relied on the instantaneous alpha frequency estimation approach which has since then been shown to be flawed.

4. Alpha power is another important confound. The authors show that there is not a significant difference in alpha power using classical statistics. However, there is non-decisive evidence using Bayesian analysis (in fact, even weak evidence for a power difference). Hence, alpha power confounds cannot be completely excluded.

Despite these concerns, the study may contribute to the field by sparking further research into the role of alpha frequency. However, ideally (and as previously suggested), the authors should address the methodological concerns and employ techniques known to provide unbiased alpha frequency estimates (i.e. FFT over finite timewindows).

Reviewer #3

(Remarks to the Author)
The authors have addressed all my comments and suggestions in great details. I commend the authors on the work and care invested in responding to each comment or concern I had.
I believe the manuscript was greatly improved and that this is a well-designed and very impactful study.

Reviewer #4

(Remarks to the Author)
I was invited to substitute for Reviewer 2 and assess whether the concerns raised have been appropriately addressed. First and foremost, I share the highly positive impression of the manuscript and would like to congratulate the authors on their outstanding work. The study represents an ambitious effort to scrutinize the relationship between IAF and visual perception, featuring a large sample size, multiple analytical approaches that converge on a main conclusion, and a conceptual integration of distinct research strands (perceptual and temporal sampling). Moreover, it generates testable research questions that will undoubtedly inspire future studies. Taken together, these strengths make the manuscript a strong candidate for publication in Nature Communications, which I strongly endorse.

A significant portion of Reviewer 2's comments focused on wording, typos, overlooked literature, and the need for clarifying remarks (e.g., regarding the strength of effects and the scope of the IAF-perception relationship, i.e., restriction to only very short stimuli). I believe the authors have done an excellent job addressing these concerns both in their response and in the revised manuscript.

The most critical issue raised by Reviewer 2 concerns the potential influence of alpha power on instantaneous frequency. The authors employ multiple approaches to rule out a statistically significant impact on the IAF effect, though one could argue that the statistical approach in Figure S2 is overly conservative, as it tests a very broad frequency and time range.

However, my concern lies more with the reliability of extracting IAF at the single-trial level, which is crucial for several of the presented analyses. The methods section states that individuals without a clear IAF were excluded (ll. 582–584), but this does not fully account for various ways in which the power spectrum might influence the findings. For instance, changes in non-alpha frequency bands associated with more accurate perception could “leak” into the narrow-band filtered data, thereby affecting the estimated peak frequencies.

To address this, it would be helpful to present Figure S2 in power spectral form over a narrower range (e.g., up to 20 Hz) for both conditions, overlaid with an estimate of variability. Another potential issue is that in trials where participants are less accurate, noisier data could make alpha peak estimation less reliable. One way to test this would be perhaps to use the “aperiodic-free” single-trial power spectra and fit Gaussians to estimate peak frequency and Goodness-of-Fit per trial. If the effect is genuinely driven by IAF, Goodness-of-Fit should not differ between conditions.

That said, compared to the field addressing similar questions, the authors have already conducted extensive analyses. Given that I am joining the review process at a late stage, I wouldn't insist on these additional checks but suggest them as potential refinements or discussion points.

Although addressing this issue is not strictly within my role as a substitute for Reviewer 2, I found the response to Reviewer 1's concern about the potential influence of eye movements to be overly defensive. While ICA correction removes a significant portion of instantaneous volume conduction effects, it does not necessarily eliminate the neural activity associated with eye movements. There is exciting ongoing research exploring the relationship between oculomotor behavior and alpha activity (see, for example, work by Ole Jensen's lab). If the authors wish to further investigate this, they could consider analyzing the 'blink component' (acknowledging that this primarily captures vertical movements) or defining a frontal electrode group for analysis using the non-eye-movement-corrected data. However, I do not insist on these additional analyses.

Version 2:

Reviewer comments:

Reviewer #4

(Remarks to the Author)
The authors have done a great job in addresing my concerns. This is a great paper. Congratulations.

Reviewer #5

(Remarks to the Author)
In an EEG study of the role of alpha oscillations in visual perceptioin, the authors investigate how indivdiual and instantaneous alpha frequency (IAF), together with alpha phase, influence visual accuracy and sensitivity in a visual detection task. Their key findins is that IAF and phase jointly modulate visual perception, and the authors provide data and derive a qualitative model that lower IAF makes the impact of phase more likely. The study is well designed, uses a large sample (n =125) and sophisticated analysis methods including SDT and DDMs. Overall, the work appears as an important contribution in the ongoing debate of whether IAF modulates visual perception, including a novel view how IAF may link to alpha phase effects on visua perception.
I have been added as a reviewer after first rounds of revision, and have been asked to especially focus on a previous reviewer (#1) comment regarding the computation of instantaneous IAF, potentially being biased by 1/f noise and SNR, power confounds, and eye motion confounds. So I mainly focus on the author's reply to these concerns:
• Inst. IAF approach: The authors provided four additional analyses where they avoided computing instantanous IAF, but computed and compared IAF from power spectra of trials with correct vs. incorrect response using establised methods (Corcoran protocol, FFT methods in a fixed prestimulus interval). Specifically, they could reproduce their main findings (higher IAF has better visual accuracy/sensitivity) when estimating alpha peaks from single-trial data using the Corcoran approach. Further, the authors provide an alternative approach to remove 1/f noise using polynomial fitting. In all cases, the author could replicate their main findings. While I share the concerns of reviewer #1 that the instantaneous IAF may be partially a misnomer because frequency cannot be estimated instantenously (i.e., it needs extensive time windows for computation to achive sufficient frequency resolution to distinguish tiny frequency differences between conditions), the author provide convincing complementary analysis their their results are robust using other methods, avoiding the inst. IAF method's potential pitfalls.
• Power confounds: The authors provide additional analyses showing that power differences between correct and incorrect decision are absent or minimal (also using Bayes factors), esp. compared to the IAF effects. (Minor: the supplemental fig. seems to lack a color bar description and a figure legend). So power confounds of the IAF effects appear unlikely.
• Previous reviewers also mentioned potential eye motion confounds, which were adressed previously. I do not think that eye motion confounds are specifically problematic in this study because at prestimulus time points, before any stimulus, participants have no information to alter their eye motion behavior. In principle, a link between differentiial eye motion (e.g. blinks) and IAF is conceivable, where IAF then arose from the eye motion, but this appears rather unlikely. A simple control analysis would be to compare e.g. blink frequency before stimulus onset between correct and incorrect trials.

Overall, I want to emphasize that the study is very well conducted and an important contribution. When reading the manuscripts, some additional points (none major) came to my attention, that may help to further improve the work:
Intro:
• The theoretical link between IAF and alpha phase appears plausible on first sight, because faster IAF leads to faster cycling trough optimal and non-optimal phases, so that the optimal phase is more likely hit by the stimulus. On the other hand, faster cycling also means that the oszillator stays only for a shorter period within the optimal phase, compared to a slower IAF. So the 'positive' effect of faster phase cycling could be well offset by a shorter duration within the optimal phase. This could be formalized/visualized by a phase-over-time plot.

Results:
• Time-resolved analyses: Strikingly, sensitivity and accuracy already diverge (even though non-significantly) much earlier than the stimulus onset, already at -800ms. One would expect that at least for very early prestimulus time points (e.g. -1.5 s),

IAF is not predictive of sensitivity/accuracy. Analyzing such a temporal emergence over longer time periods would serve as an important control, enhancing the specifity of a main finding.
• Alpha phase analysis: The author report alpha phase effects that occur around stimulus onset (starting at -49ms) or after stimulus onset. Thus, it appears that the phase effect is related to stimulus processing, because the prestimulus effect could well arise from temporal smearing of the time-frequency analyses into the the peri-stimulus window. So the phase effect may not arise from ongoing alpha oscillations. Further, the authors may show a (supplemental) figure of the time course of the effect.
• IAF x phase interaction analysis: This analysis mixes within- (phase) and between (IAF) subject levels: Showing that a phase-effect exists in low IAF individuals, but not in high IAF individuals does not allow the conclusion of a significant difference of phase effects between IAF groups (i.e. absence of evidence is not evidence of absence). This could be formally assessed with an interaction effect in this analysis. In principle, this analysis could also be computed purely within individuals, e.g. using a 2 x 2 binning approach on IAF x phase bin, showing an interaction effect. The trial-by-trial analysis of interaction effects shows that this is in principle the case, but a 2x2 approach could well visualize the effect.
• ITPC analysis: Again, the question is whether correct-incorrect contrast differs between high and low IAFs. This interaction effect is not tested directly.
Discussion:
• Fig. 7: Even though the model is convincing at first sight, it makes a critical assumption: For accurate visual perception, the optimal phase must be hit, independent for how long stimulus processing falls into this phase. Under this assumption, higher IAF will indeed increase the likelihood that a stimulus at least briefly touches the optimal phase due to rapid cycling. On the other hand, the amount of time a stimulus is processed during the optimal phase could also be critical for accurate perception. Here, a lower IAF may increase the probability for a high optimal-phase duration. Thus, these two effects might cancel each other. Only a quantitative model could distinguish these cases.
Minor
• Line 335: The polar plot should only be presented describtively, but phases should not be tested again because they arise from the already significant cluster (double dipping).
• Line 434: Reference to Fig. 7 may make more sense?

Version 3:

Reviewer comments:

Reviewer #5

(Remarks to the Author)
The authors responded extensively and convincingly to my previous comments, adding some important specifications to their model and adding important control analyses:
• As stated in my previous comments, I find the control analyses in response to reviewer #1 comments convincing.
• The authors refined their model to link IAF and phase, accounting now for the possibility that duration within an optimal phase may play a role. I agree with their view that rapid perceptual sampling is more important for the system in dynamic environment than the sampling duration of a stimulus in the optimal phase – sensory systems need to represent change, not constancy. On the other hand, if only the speed of cycling is crucial, why does the brain use a comparably slow oscillation around 10 Hz for perceptual sampling, and not, say, 100 Hz? There must be some upper limit in the perceptual sampling frequency, which arise from internal (e.g. “over-sampling” leads to unnecessary information load) or external factors (e.g. the temporal statistics of usual environment do not require fast sampling).
• Further, the authors performed time-resolved analyses of the predictive IAF effect on perceptual performance that extended the time window to much earlier time points. This analysis shows that the effect emerges much early around -1.5s prestimulus, but not at even earlier time points. This is an important control analysis and finding.
• Finally, the authors convincingly argue and provide a control analyses showing that a temporal smearing explanation of the phase effect shortly before stim onset is unlikely.
• Finally the author performed a formal test of the interaction effect of IAF x phase as suggested. This analysis confirmed an interaction shortly before stimulus onset. I agree that a similar within-subject analysis of ITPC for IAF x accuracy interaction is impossible if cells of the 2x2 design have unequal trial numbers, but the authors show with a mixed analyses that such a prestimulus interaction is indeed present in their data.

Reviewer #1 (Remarks to the Author):

The authors conclude that alpha frequency influences sensory precision by modulating the distribution over phases.

There are a couple of conceptual and metholodological issues:

We would like to thank Reviewer 1 for their insightful comments, which have enabled us to further detail the nature of our effect and clarify that the methodology employed for peak estimation is state-of-the-art, allowing for highly precise frequency resolution. Additionally, these observations have helped us contextualize our findings within the existing literature on the relationship between alpha oscillations and perceptual sensitivity. Below, we have provided a point-by-point response to all comments received.

1. Can the hypothesized mechanism explain variation in observers' behaviour?

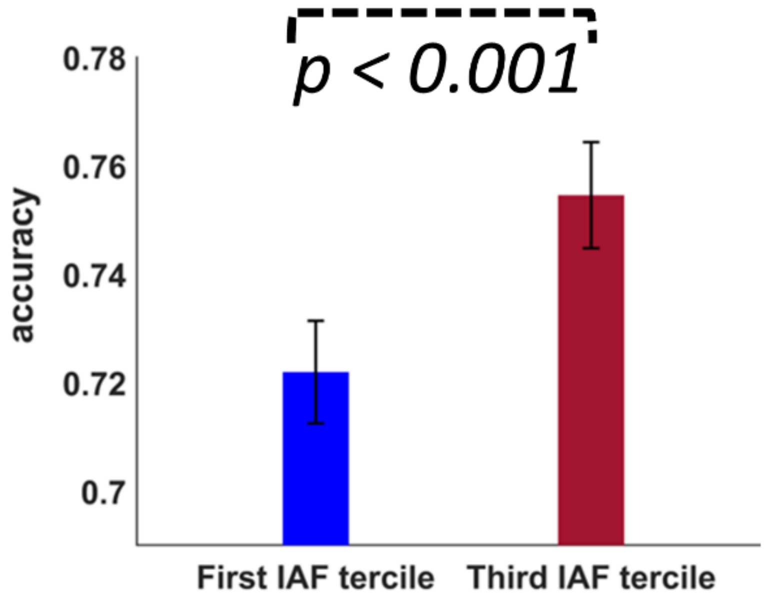
The authors suggest that higher alpha frequency increases the probability that the optimal phase will be present during stimulus presentation. Let's think this through: stimulus duration = 60ms. The authors observed an alpha frequency of 11.15 Hz (correct trials) vs. 11.12 Hz (wrong trials) i.e. a difference in alpha frequency of 0.03 Hz. Hence, the length of the alpha cycle decreases from 89.928 ms for wrong trials to 89.686 ms for correct trials. The coverage of alpha phases for wrong trials about 66.7% and for correct trials about 66.9% , i.e. a tiny increase in the coverage of alpha phases of 0.2%. Can such a tiny increase in the coverage of alpha phases really explain a measureable behavioural effect? Can it explain a behavioural effect that can be measured reliably with only 150 trials per participant?

Thank you for your comment. We believe this perspective stems from viewing the mean observer as an "actual" observer. However, it is important to clarify that the mean observer represents an arithmetic aggregate of mean values both within and across participants. Indeed, the variability reported by looking at the "mean observer" does not represent the range of variability within and between participants, and as such it hampers the impact that instead bin-based and trial-by-trial analysis provided. The mean observer effect was reported not as a main finding but as a confirmatory finding supported by strong frequentist and Bayesian effect statistics. It just means that intraindividual and interindividual variability are important mediators of this effect as shown in the set of analyses provided. Indeed, despite the intra- and inter- individual variability, the small difference noted in this analysis between correct and incorrect responses, as highlighted by Reviewer 1, does not diminish the findings. On the contrary, this result is exactly what one might expect given the nature of the phenomenon, it is therefore highly plausible and reflects previous effect sizes by seminal works (see below). Incidentally, following the improvement of the filter prompted by the reviewer's useful comment, we have demonstrated that this difference increases to 0.05 Hz, and we will refer to this value henceforth. Thus, we believe this does not undermine the plausibility of an interaction between individual alpha frequency (IAF) and phase; rather, it reinforces the replicability and robustness of these findings, which have been similarly reported by several independent laboratories. Below we present 5 key points to support this statement:

1) **Within-Individual Variability**: The observed difference in IAF between correct and incorrect trials might be small at the arithmetical level because the phase coverage is quite similar within the same individual. To illustrate this, we calculated the standard deviation of IAF across time, trials, and participants, resulting in a value of 0.89 Hz (strikingly close to the value of 0.9Hz highlighted by the seminal work of Haegens et al., NeuroImage, 2014). Therefore, for a participant with an IAF of 10 Hz, the range of IAF values will fall between 8.2 and 11.8 Hz in 95% of cases (MEAN ± 2 STD). Comparing these two extreme cases, the stimulus coverage would be 49% for an IAF of 8.2 Hz and 71% for an IAF of 11.8 Hz. This important difference in coverage is in line with the hypothesis that a higher IAF significantly increases the likelihood of accurate responses compared to a lower IAF due to this differential phase coverage. However, in other scenarios where the within-subject IAF difference is less pronounced between different trials, the coverage would show less variation, as the reviewer correctly pointed out. For example, with a 0.2 standard deviation difference (10 Hz - 0.2 * 0.89 = 9.82 Hz; 10 Hz + 0.2 * 0.89 = 10.18 Hz), the coverage would be 58.9% for an IAF of 9.82 Hz and 61.1% for an IAF of 10.18 Hz. Therefore, in this case, the likelihood of the stimulus falling within the optimal phase is similar.

Connected to this point, it has to be considered the strong impact of intra-participant averaging procedure we employed. When averaging IAF within each participant across all time points (800 points) and trials (150 trials), this process smooths out inter-time and inter-trials fluctuations, reducing the variability of IAF values. Differences that might be significant at the single trial and timepoint become less pronounced when averaged across numerous points and trials. This step alone reduces the impact of any extreme IAF values by bringing them closer to the participant's mean IAF.

Crucially, we have evidence validating this hypothesized scenario empirically. In fact, recognizing the significant impact of the averaging procedure, we incorporated a bin-based analysis in the initial version of the manuscript (see image attached below). This analysis divided trials into the 1st and 3rd bins based on individual alpha frequency (IAF) values on a subject-by-subject basis. Crucially, this analysis reduces the smoothing effect of averaging at the intra-individual level, allowing us to consider the trials in which the deviation from the mean was consistent (on average, above (3rd bin) vs. below (1st bin) 1 SD from the mean). This is proved by the simple observation that in the first bin, the IAF average across participants was 10.90 Hz, while in the third bin, it was 12.19 Hz. So, by smoothing out the effect of the mean, the difference is around 1.3 Hz between the first and the third terciles. This value is about 26 times larger than the difference of 0.05 Hz. In this case, on average, in the first bin the alpha cycle covers approximately 65.4% of the stimulus, while in the 3rd bin the alpha cycle covers approximately 73.2% of the stimulus. Considering 116 individuals with 150 trials each, the difference created a notable impact. Indeed, the difference in phase coverage is compatible with the highlighted behavioral effect, in which the average accuracy in the third bin is 3 points higher than the accuracy measured in the first bin.



2) **Moreover, another crucial step to consider is the impact of inter-participant averaging**. After averaging IAF values within each participant, these averaged values have been further averaged across all 116 participants. This inter-participant averaging further reduces variability. For all these reasons, differences in IAF between correct and incorrect responses that were apparent within individuals are less likely to stand out in the grand average. The averaging process integrates a wide range of individual IAF differences, encompassing varying degrees of response accuracy and inherent IAF variability. Consequently, the differences between correct and incorrect responses observed at the individual level tend to converge towards the mean when averaged over a large number of participants. **Indeed, this process exemplifies regression toward the mean**, a statistical phenomenon where extreme values tend to move closer to the average upon repeated measurements or averaging. By averaging IAF values within and across participants, extreme differences are moderated. As a result, the final estimate shows smaller differences in IAF between correct and incorrect responses than might be seen in individual trials. Consequently, the overall IAF differences are less pronounced in this type of analysis, as the repeated averaging process obscures the more extreme values and highlights those closer to the mean IAF. This is precisely the reason that led us to also adopt trial-by-trial analysis that are able to weigh the single-trials impact, avoiding averaging procedure (indeed, this procedure

allowed us to find an even greater Bayes Factor in the trial-by-trial variability that was however already high in the mean value analysis).

Moreover, in acknowledging the influence of inter-individual difference, we incorporated an additional analysis in the initial version of the paper that we find compelling in illustrating the relationship between IAF and phase. Specifically, we replicated the binning-based approach, this time binning alpha frequencies not within participants but based on their inter-individual IAF. Specifically, we separated participants into two groups: one with high alpha frequency (average 12.38 Hz) and another with low alpha frequency (average 10.29 Hz). Once again, by avoiding averaging across the entire sample, the differentiation between individuals becomes evident — a difference of 2.09 Hz, which is approximately 42 times greater than 0.05 Hz.

We hypothesized that individuals with higher IAF would sufficiently cover the stimulus, regardless of the specific phase angle at stimulus onset. Faster alpha frequencies naturally provide greater coverage, allowing individuals to adapt even if the stimulus begins in an unfavourable phase. This is also supported by the greater variability characterizing individuals with faster IAF, as empirically demonstrated in the revised manuscript (Fig. 1C). Therefore, if variations in alpha are unable to bring significant differences in the covered phase angle, the role of phase should be less crucial. Conversely, we expected individuals with slower IAF to be more phase-dependent, unable to effectively compensate for starting from an unfavorable phase angle. Our results exactly confirmed this hypothesis: individuals with higher alpha frequencies showed a lesser impact of phase on performance, while those with lower alpha frequencies were significantly influenced by phase. The difference in frequency between these groups (i.e., 2.09 Hz) supports this disparity. For example, individuals with faster alpha cycles (80.77 ms) would cover approximately 74% of an alpha cycle within the 60 ms stimulus window, whereas those with slower alpha cycles (97.18 ms) would cover about 62%. Therefore, even though there is inter-trial and inter-time variability as described in point 1, this has less impact on individuals with a high IAF because, even in the "worst" cases (i.e., 2DS below the mean), a high IAF still allows for sufficient coverage of the stimulus. So, the magnitude of the different phase coverage between the two sub-populations is significant and thus explains the interaction IAF*Phase we robustly demonstrated. In summary, we repeat again that, while the observed differences in IAF may appear small at the grand average level, they are meaningful when considering the broader distribution and variability within individuals. These differences, even if minor in their mean absolute value, can significantly influence behavioral outcomes, supporting the conclusion that alpha frequency influences perceptual sensitivity by modulating the distribution over phases.

3) To provide a broader perspective beyond our own research group, we present 3 figures from highly relevant studies published in top-tier journals (including one in Nature Communications) investigating the relationship between IAF and perceptual outcomes (Nelli et al., Nat. Comm. 2017; Samaha et al., Curr. Biol., 2015; Wutz et al., PNAS, 2018). As it can be clearly seen, the magnitude of differences reported between correct and incorrect responses strictly mirrors what is reported in our findings. This underscores that even subtle differences, which appear subtle only in terms of their mean absolute value for the reasons discussed earlier and which apply equally in these studies, significantly influence perception when averaged across numerous trials and participants.

Nelli et al., Nat Comm, 2017
https://doi.org/10.1038/s41467-017-02176-x

Figure Redacted

Samaha et al., Curr Biol, 2015

https://doi.org/10.1016/j.cub.2015.10.007

Figure Redacted

Wutz et al., PNAS, 2018

https://doi.org/10.1073/pnas.1713318115

Figure Redacted

4) Another very relevant point concerns the statistical interpretation of the findings. A fundamental principle in statistics is that relying on arithmetic or absolute values to assess significance and impact can be misleading. This approach is similar to basing the presence of significance on a figure that represents the average of two measurements. It is important to note that both the analysis highlighted contrasting IAF in correct vs. incorrect trials and the bin-based analysis are statistically compelling, as shown by the Bayes Factor above 50, which strongly supports the hypothesis linking IAF to accuracy.

5) Regarding the possibility of detecting this effect with just 150 trials per participant, we would like to point out that, while it is true that we have approximately this number of trials per participants, we tested **125 participants** [116 included in the final analyses after removing 9 as prescribed by the guidelines indicated by (Samaha and Cohen, 2022)], which we believe is the largest sample size ever used to investigate the relationship between IAF and perception. To put this into perspective, consider that a recent paper that has attracted significant attention by Buerges and Noppeney (Nat. Hum. Beh., 2022), which concluded no relationship

between IAF and perceptual sensitivity, included only 20 participants—a huge contrast to our significantly larger sample size (more than 100 additional participants). Furthermore, even if one considers that 150 trials may introduce a low signal-to-noise ratio by increasing the noise component due to their relatively limited number, our observed results demonstrate that the effect is substantial enough to effectively overcome any potential noise introduced.

In conclusion, we hope the reviewer will now convene with us on the methodological and statistical strength of our observed effects which indicate their robustness and significance in the relationship between IAF and perceptual sensitivity.

2. Is it possible to measure a change in alpha frequency of 0.03 Hz reliably given the constraints imposed by the experiment?

To estimate a 0.03 Hz difference reliably and obtain such an extremely high spectral resolution, one requires very long data segments. However, the segments are limited to only brief 700 ms prestimulus baseline.

Moreover, the estimation of the instantaneous phase with the Hilbert transform depends on the filter parameters. No data are shown that confirm the appropriateness of a filter range of IAF +/- 2 Hz.

The authors take the first temporal derivative to remove an aperiodic component. However, this is often not sufficient to remove 1/f noise as an important confound.

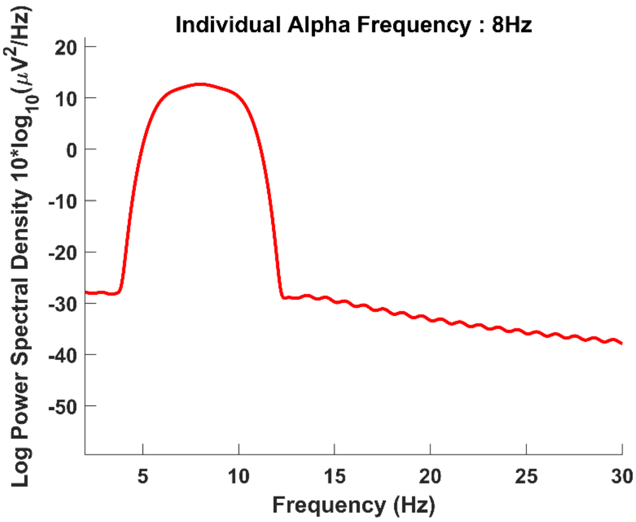
We completely agree that detecting a difference of 0.03 Hz (as well as 0.05Hz) requires extremely high spectral resolution. However, to avoid the limitations of classic FFT analysis, which would have hindered our ability to detect this subtle difference, we utilized instantaneous frequency analysis to overcome this constraint (Cohen, MIT press, 2014). Importantly, the vast majority of studies published in top-tier journals investigating the role of IAF in perception (Benwell et al., NeuroImage, 2022; Buccellato et al., NeuroImage, 2023; Buerges and Noppeney, Nat. Hum. Beh., 2022; Drewes et al., Cerebral Cortex, 2022; Nelli et al., Nat. Comm. 2017; Samaha et al., Curr. Biol., 2015; Sharp et al., J. Neuroscience, 2022; Shen et al., Plos Biology, 2019; Wutz et al., PNAS, 2018) have utilized the very same data-analysis approach specifically to detect these subtle effects (see previously attached figures). Furthermore, we have employed an even more robust approach than previous papers by individually selecting the filter, centered around the IAF, on a subject-by-subject basis, as recommended by relevant methodological work on this topic (Samaha and Cohen, NeuroImage, 2022).

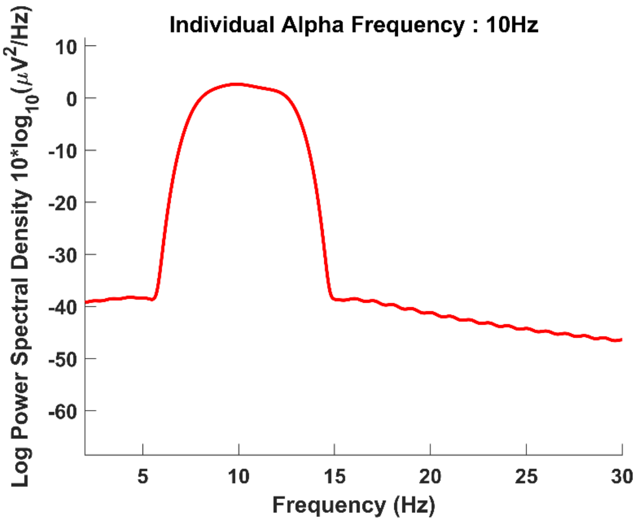
Moreover, as previously shown (see previously attached figures), the average difference observed in our study (0.05 Hz) dramatically aligns with findings from these influential and pivotal studies examining differences between correct and incorrect responses using different experimental protocols. Therefore, suggesting that such differences cannot be reliably traced implies fundamental flaws in both our study and these seminal studies. If such flaws existed, it would necessitate a significant reassessment of the field. However, we believe that asserting that the effect cannot be measured is a very strong and unjustified claim, potentially undermining a decade of research that has been corroborated and widely replicated.

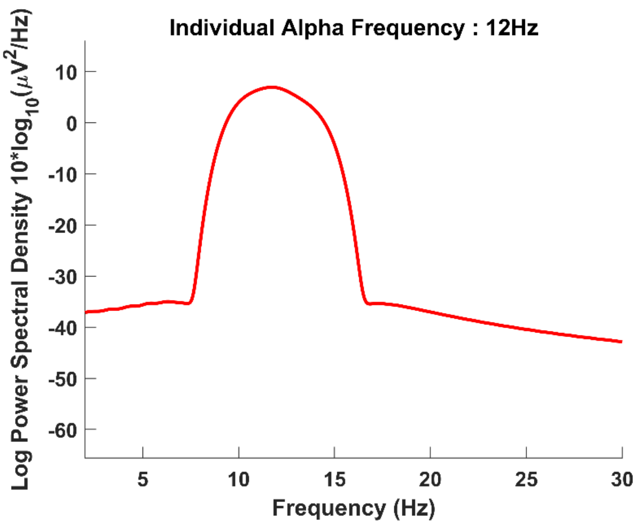
Please note that our study, however, went well over replicating this mean average effect, instructing on the mechanisms determining the how alpha frequency impacts sensory sampling.

Last but not least, related to the previous points, this relationship is notably robust, as supported by the Bayes Factor (BF) being above 30 across the majority of analysis (again, indicating strong evidence in favor of H1). We believe that it is very unlikely that such a strong effect arose purely by chance, if it cannot even be measured properly.

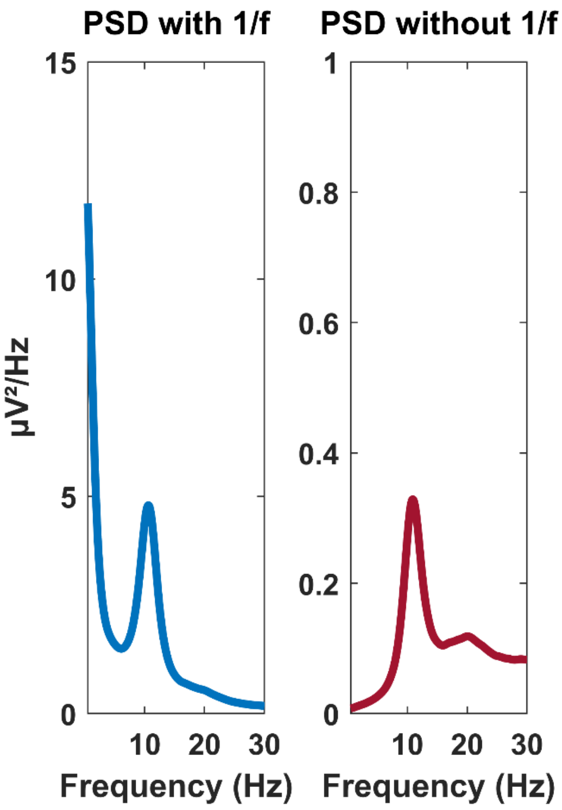
Moreover, we agree with the reviewer about the importance of designing an appropriate filter for extracting the instantaneous alpha frequency. Upon review, we found that the initial filter employed slightly different parameters compared to those commonly used in the literature (London et al., 2022), so we adjusted it accordingly. Specifically, we filtered the signal using a zero-phase, plateau-shaped, band-pass filter with 15% transition zones (Matlab function *filtfilt.m*). It is important to highlight that the overall pattern of results remains fully consistent with the initial version. Attached are some figures showing the power spectrum obtained using the *spectopo* function after the filtering process for selected subjects. As observed, the filter effectively captures the alpha band and crucially adapts to the precise range based on each subject's individual alpha frequency, while frequencies outside this range are attenuated.


Individual Alpha Frequency : 8Hz
Log Power Spectral Density 10*log10(μV²/Hz)
20
10
0
-10
-20
-30
-40
-50
5
10
15
20
25
30
Frequency (Hz)


Individual Alpha Frequency : 10Hz
Log Power Spectral Density 10*log10(μV²/Hz)
10
0
-10
-20
-30
-40
-50
-60
5
10
15
20
25
30
Frequency (Hz)


Individual Alpha Frequency : 12Hz
Log Power Spectral Density 10*log10(μV²/Hz)
10
0
-10
-20
-30
-40
-50
-60
5
10
15
20
25
30
Frequency (Hz)

Finally, we are happy to present a comparison of the power spectrum averaged across subjects (see attached figure), both with and without the removal of the 1/f aperiodic component. As clearly shown, our technique effectively isolates and removes the aperiodic noise, while preserving the essential periodic features of the signal. A key indicator of this is the alpha peak, which remains relatively consistent between the two plots, demonstrating that alpha activity was effectively preserved throughout the procedure. Further evidence supporting the efficacy of our approach is observed in the beta band: in the left plot (without 1/f removal), no distinct peak can be extracted, whereas the right plot (with 1/f removal) reveals a prominent beta peak, clearly visible due to the successful attenuation of aperiodic noise.



1. Changes in power which are a key confound are not reported.

We noted that not only Reviewer 1, but also Reviewers 2 and 3 requested these control analyses on alpha power, and we fully agree that their inclusion would strengthen the proposed framework.

We fully agree with them on the importance of clarifying the specificity of the effects linking IAF to perceptual performance, ruling out the possibility that the same relationship could be driven by alpha power. To address this concern, we conducted a series of analyses that consistently demonstrated that alpha power does not significantly influence perceptual performance. Below is a summary of the analyses, which are detailed in the supplementary materials.

First, we extracted time-frequency amplitude maps and conducted a cluster-based permutation test to assess whether oscillatory amplitude differed based on response accuracy (correct vs. incorrect trials). The statistical analysis revealed no significant clusters distinguishing correct from incorrect responses ($p = 0.43$).

Next, we performed regression analyses using a least-squares approach, adding alpha power as an additional predictor alongside IAF. The results indicated that IAF remained a strong predictor of performance (slope = 0.37, SEM = 0.09; $t_{115} = 4.04$, $p < 0.001$, BF = 179.63), while alpha power (mean slope = -0.14, SEM = 0.09; $t_{115} = -1.57$, $p = 0.12$, BF = 0.34) did not demonstrate any predictive value. A similar pattern emerged when using a GLM model (mean slope for alpha power = -0.03, $p = 0.13$; mean slope for IAF = 0.08, $p < 0.001$).

We also conducted bin analyses by dividing trials into terciles based on alpha amplitude. The results showed no significant differences in sensitivity and criterion between trials in the first and third terciles (all $p > 0.09$, all BF < 0.41).

Finally, we conducted another time-frequency analysis comparing oscillatory amplitude between participants in the slow and fast IAF groups. No significant clusters were found between the two groups ($p = 0.14$).

In conclusion, these analyses consistently demonstrate that alpha power does not contribute to explaining perceptual performance, in contrast to the robust influence of IAF.

Here we report the new section added in the methods section:

Assessing the specificity of the relationship between IAF and Perceptual Accuracy relative to Alpha Power

To ensure that the effects observed in relation to IAF were not confounded by oscillatory power, we conducted several analyses (See Supplementary Materials). First, we extracted time-frequency oscillatory amplitude maps for each participant from the specific electrodes used to compute IAF. Time-frequency decompositions were performed for 50 linearly spaced frequencies ranging from 2 Hz to 50 Hz. Wavelets with 3 cycles at the lowest frequency and 11 cycles at the highest frequency were applied. After obtaining the absolute value of the resultant analytic signal, we conducted a cluster-based permutation test to evaluate whether oscillatory amplitude varied based on response accuracy (correct vs. incorrect trials). Next, we extracted the trial-by-trial amplitude from the time-frequency analysis, averaging it within the -800 ms to -100 ms window, at the frequency closest to each participant's previously determined IAF. We performed regression analyses using a least-squares approach, incorporating both trial-by-trial fluctuation in IAF and alpha power as predictors of perceptual accuracy. This analysis allowed us to determine whether the influence of IAF on performance was independent of alpha power. We also replicated a bin-based analysis by dividing trials into terciles based on the previously extracted trial-by-trial alpha amplitude. We compared signal detection theory metrics across the lowest and highest terciles to confirm that the effects observed with IAF did not hold when considering variations in alpha power. Lastly, we conducted a time-frequency analysis using the same parameters as those applied in the comparison of correct and incorrect responses. This analysis aimed to assess potential differences in oscillatory amplitude between participants categorized into low and high IAF groups.

Here we include the new part and figure included in the supplementary materials:

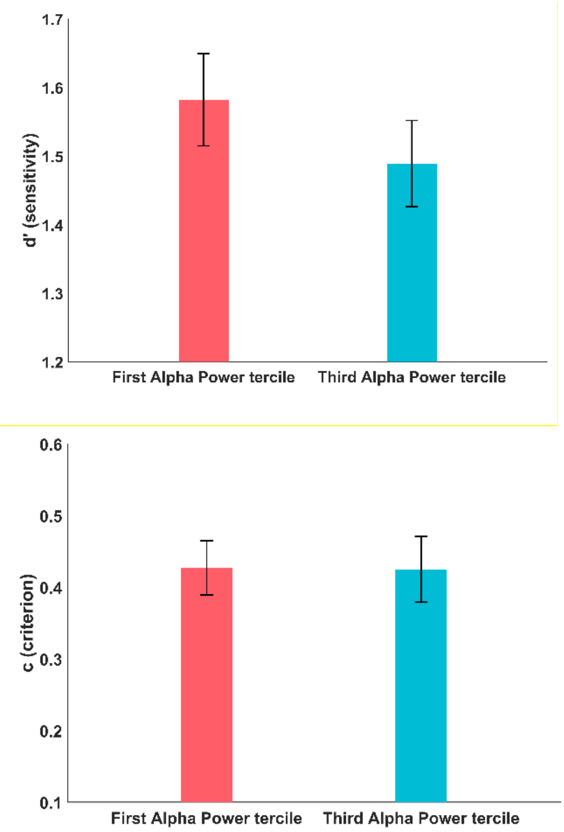


Figure S1. No significant differences observed in behavioral performance between first and third power terciles

We replicated the initial analysis included in the main text by binning trials based on pre-stimulus power in the first and third terciles. The analysis revealed that this binning approach did not dictate behavioural performance (all $t_{115} < 1.68$, all $p > 0.09$, all BF < 0.41).

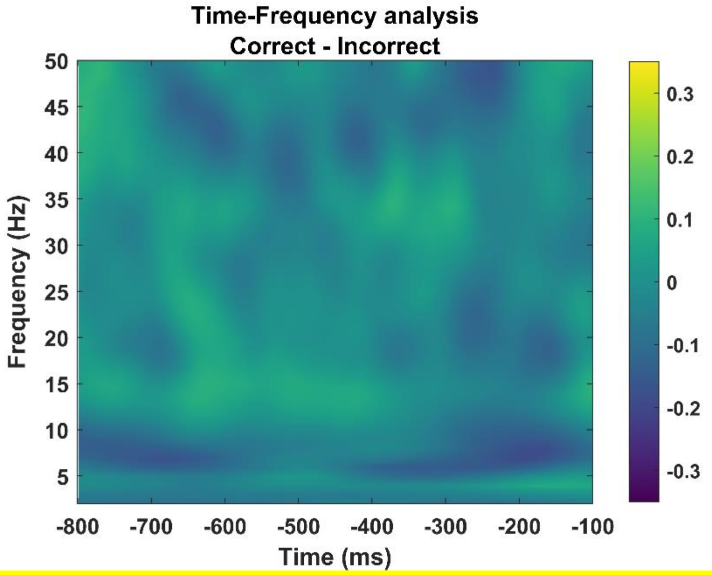


Figure S2. Correct and Incorrect trials are not characterized by difference in oscillatory amplitude.

We conducted a time-frequency analysis to assess whether correct responses exhibited different amplitudes compared to incorrect responses. The cluster-based analysis revealed no significant differences between the two types of trials, indicating that power, in contrast to IAF, is not capable of differentiating correctness in the trials.

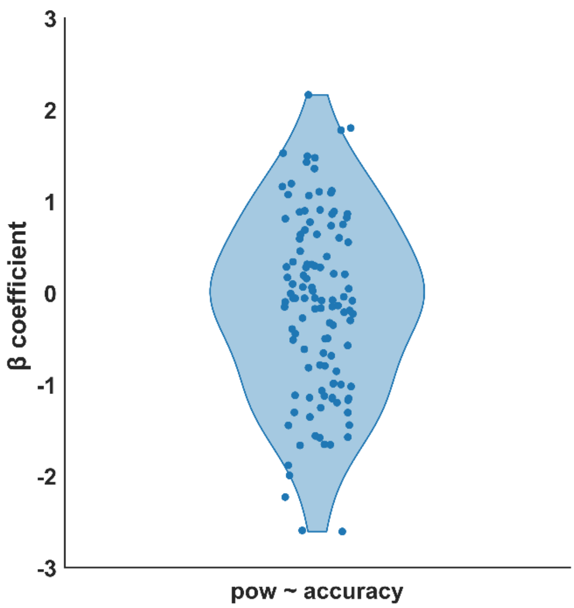


Figure S3. Trial-by-trial fluctuations in alpha power do not account for perceptual accuracy.

We conducted an analysis to evaluate whether trial-by-trial fluctuations in power, alongside those of IAF, influenced the accuracy of the responses provided. We extracted the power calculated from the time-frequency analysis by taking the alpha frequency closest to the IAF, normalizing it through a z-score transformation. Subsequently, we created a design matrix that included the intercept, IAF, and normalized alpha power. The regression analysis confirmed the significant predictive ability of IAF fluctuations on perceptual accuracy (mean slope = 0.37 $p < 0.001$). In contrast, fluctuations in power did not modulate the participant's ability to accurately retain the stimulus (mean slope = -0.15 $p = 0.12$).

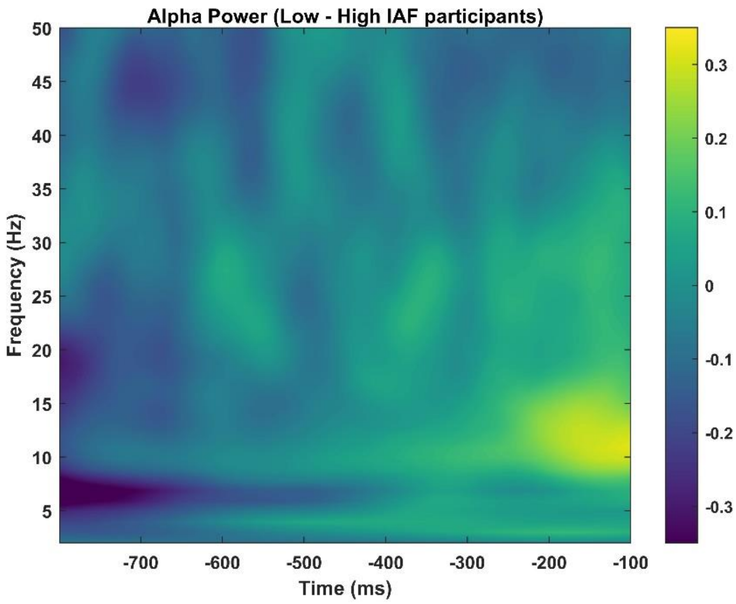


Figure S4. Low and Fast IAF groups do not exhibit significant differences in oscillatory amplitude.

We conducted a time-frequency analysis to assess whether participants in the low IAF frequency group exhibited different amplitude profiles compared to those in the high IAF frequency group. The cluster-based analysis revealed no significant differences between the two populations, indicating that power does not effectively differentiate between the two groups.

2. No eye tracking data are reported. Eye movements are crucial confounds that need to be excluded.

We agree that eye-related artifacts could potentially act as a significant confound in the EEG analysis we conducted. However, it is important to note that we have already addressed potential confounds related to eye movements by employing Independent Component Analysis (ICA), which is a widely recognized technique for offline artifact removal, in the first version of the manuscript. This method is the gold-standard to eliminate artifacts associated with eye movements from our data. Finally, it is noteworthy that the seminal studies described in the above comments (Nelli et al., Nat. Comm. 2017; Samaha et al., Curr. Biol., 2015; Wutz et al., PNAS, 2018) also employed the same artifact removal techniques (i.e., ICA) in order to remove this type of artifact.

3. The authors report only results at the sensor level. This makes it impossible to draw decisive conclusions about the underlying neural mechanisms. It is very likely that alpha frequency as measured on the scalp arises from multiple neural generators e.g. parietal and occipital sources which may have different frequencies. Hence, putative changes in alpha frequency may reflect changes in power of different alpha generators (rather than a change in alpha frequency).

This is a great comment. We agree that addressing the possibility of spurious effects resulting from an artificial "merging" of alpha activity generated by different sources would be a valuable enhancement to the paper. In response to the Reviewer observations, we would like to clarify that we have conducted source-level analyses that provided more detailed insights into the underlying neural mechanisms.
Our source-level analysis results clearly demonstrate that alpha activity originating from the occipital lobe serves as a strong predictor of the observed behavior. In contrast, while alpha activity generated from the right parietal lobe shows some significant associations with behavioral measures, these associations do not withstand a Bayesian approach, which reveals a higher probability for the null hypothesis (H0) compared to the alternative hypothesis (H1) in many cases. This suggests that the influence of the parietal lobe may be less robust, further supporting the notion that the association between interindividual alpha frequency (IAF) and perceptual sensitivity is more directly related to occipital activity. We believe these source-level results not only clarify the roles of different brain regions in alpha activity but also provide a deeper understanding of the neural mechanisms involved in the studied behavior. It is important to highlight that, given the limited spatial resolution of EEG, we cannot definitively conclude that our occipital and parietal signals were fully separated; there may have been residual mixing of signals even after source reconstruction. Nevertheless, we observed a divergence between occipital and parietal

alpha peak frequencies in shaping perceptual performance. Overall, we are confident that this integration of source-level analyses strengthens the conclusions of our work.

Specifically, we have added in the main text that:

> *we also ruled out the hypothesis that the relationship between IAF and perceptual sensitivity is merely due to spurious changes in alpha frequency reflecting variations in the power of different alpha generators. Our source analysis demonstrates that all observed effects are robustly predicted by the IAF extracted from occipital sources.*

And we have added in SI the following analyses:

> *We addressed the concern that the observed relationship between IAF and behavior might be an artifact of merging alpha activity originating from distinct sources by performing a source-level analysis, inspired by the approach in Haegens et al. (2014). Specifically, we used the Desikan-Killiany atlas to select occipital (Pericalcarine sulcus and lateral occipital cortex) and parietal regions (Inferior Parietal and Superior Parietal cortex). For each region, we identified the sources with the highest alpha power and extracted the IAF separately for occipital and parietal sources, thereby mirroring the procedure used in our sensor-level analysis. Then, we investigated the influence of pre-stimulus IAF on perceptual sensitivity and bias by employing a binning analysis using the IAF extracted for each individual and separately for the two sources (i.e., occipital and parietal). This analysis involved dividing the trials into three terciles based on pre-stimulus IAF. Firstly, we time-collapsed the IAF data by computing the average values from -800ms to -100ms for each participant. Subsequently, we computed the signal detection indices from trials within the first and third terciles. The results demonstrated that trials in the first tercile were associated with reduced sensitivity compared to trials within the third tercile both in the right occipital (mean $d'_{first\ tercile} = 1.42$, $SE = 0.08$; mean $d'_{third\ tercile} = 1.61$, $SE = 0.06$; $t_{101} = -3.56$, $p < 0.01$; $BF = 37.44$; Fig. 1A) and in the right parietal sources (mean $d'_{first\ tercile} = 1.46$, $SE = 0.08$; mean $d'_{third\ tercile} = 1.61$, $SE = 0.06$; $t_{106} = -2.33$, $p = 0.02$; $BF = 1.43$).*

Similarly, accuracy was higher in the third compared to the first terciles when considering the occipital (mean accuracy $_{first\ tercile}$ = 0.71, SE = 0.01; mean accuracy $_{third\ tercile}$ = 0.75, SE = 0.01; t_{101} = -3.95, p < 0.01; BF = 129.50) as well as parietal sources (mean accuracy $_{first\ tercile}$ = 0.72, SE = 0.01; mean accuracy $_{third\ tercile}$ = 0.75, SE = 0.01; t_{106} = -2.78, p < 0.01; BF = 4.11). Conversely, the criterion indices estimated in the first vs. third terciles were not statistically different (Occipital source: mean c $_{first\ tercile}$ = 0.42, SE = 0.05; mean c $_{third\ tercile}$ = 0.40, SE = 0.05; t_{101} = 0.94, p = 0.35; BF = 0.17; Parietal source: mean c $_{first\ tercile}$ = 0.44, SE = 0.05, mean c $_{third\ tercile}$ = 0.43, SE = 0.04, t_{106} = 0.25, p = 0.80, BF = 0.11).

It is important to note that the strength of the relationship between IAF and perceptual performance is significantly more pronounced when considering occipital IAF compared to parietal IAF, suggesting that the alpha activity generated in the occipital lobe has a more critical impact on performance. This aspect becomes even more pronounced when we examined the differentiation of IAF in correct and incorrect responses, as well as the trial-by-trial analyses linking accuracy fluctuations to IAF variations. In this context, only the IAF extracted from the occipital source reached full statistical significance when comparing IAF in correct ($IAF_{correct}$ = 11.27, SE = 0.1) and in incorrect trials ($IAF_{incorrect}$ = 11.24, SE = 0.11, t_{101} = 2.48, p = 0.01, BF = 2.02) as well as in the trial-by-trial analysis (mean slope = 0.31, SE = 0.1, t_{101} = 3.22¸ p < 0.01, BF = 13.25), while the IAF extracted at the parietal level showed only non-significant trends, further weakened by Bayesian factors indicating a higher probability for the null compared to the alternative hypothesis ($IAF_{correct}$ = 11.24, SE = 0.1; $IAF_{incorrect}$ = 11.21, SE = 0.11, t_{106} = 1.86, p = 0.07, BF = 0.56; mean slope = 0.20, SE = 0.1, t_{106} = 1.89¸ p = 0.06, BF = 0.60).

4. The experimental paradigm involves crowding, detection etc. – possible effects of alpha frequency may act through various mechanisms making strong conclusions difficult.

All potential confounding effects were carefully considered during stimulus conceptualization and calibration. Furthermore, any potential confounds (e.g., crowding) would be consistent across conditions and individuals, as stimuli vary solely by their contrast, which is the only parameter that could explain the results. Therefore, there are no random variations that could explain the observed effect.

5. It seems that there are only 150 trials per participant. How many trials fell into different categories (e.g. correct vs. wrong)? What is the impact of the correction procedure applied to the d-prime (for small number of trials, biases can be substantial)

As reported in the paper, the average detection accuracy was 73%. Therefore, on average, 41 trials were incorrect, which are a fair number to extract alpha-related measures. Nevertheless, we implemented a subsampling procedure, matching the number of correct and incorrect responses at the individual level, in order to demonstrate that the trial count does not influence the outcome. Specifically, we included this analysis in the main text:

> *In order to check that different trials count between correct and incorrect response inflated the difference between IAF extracted from correct and incorrect responses, we used a subsampling procedure to match the number of correct and incorrect trials. Specifically, for each participant, we shuffled the trials with correct responses and matched them with IAF values count from the incorrect responses. For each permutation, we calculated the average IAF for the correct responses. This permutation procedure was repeated 500 times to ensure reliability. We then averaged the surrogate IAF values obtained from these permutations. Finally, we compared the average IAF values for correct versus incorrect responses. This comparison corroborated the initial analysis presented in the manuscript, where all trials were included as correct trials were associated with higher IAF values ($IAF_{correct} = 11.35 \pm 0.01$) compared to incorrect trials ($IAF_{incorrect} = 11.30 \pm 0.01$, $t_{115} = 4.15$, $p < 0.01$).*

Furthermore, we calculated the Signal Detection Theory (SDT) index d' and criterion by employing a different corrections strategy to ensure that the loglinear approach has not impacted the obtained results. To this end, we have used the classical approach of adjusting only the extreme rates of hit and false alarm. Rates of 0 were replaced with

0.5/n, and rates of 1 were replaced with (n - 0.5)/n, where n is the number of signal or noise trials (Macmillan & Kaplan, 1985). In the new version of the paper, we have now employed this correction method (described in the methods, see below) and changed the text and statistics accordingly.

> *To ensure a proper estimation of d' and c, we corrected the extreme rates of hit and false alarm. Rates of 0 were replaced with 0.5/n, and rates of 1 were replaced with (n - 0.5)/n, where n is the number of signal or noise trials* (Macmillan and Kaplan, 1985)

6. Quite a few methodological details are missing e.g. visual angle, luminance, stimulus duration = 60 ms with a refresh rate of 85Hz - shouldn't that be a multiple integer? Parameters for Bayesian t-test? Statistical analysis of accuracy? Generalized linear mixed effects models are recommended. Fitting procedures and parameter recovery of HDDM?

Thank you for the comment. We have added this information in the new version of the paper.

1. Regarding the information about the presented stimuli, we have added the following information in the methods section:

   Stimuli had spatial frequency of 5.16 cycles/degree, and they were presented only in the lower part of the screen at 4.1°/3.7° eccentricity (horizontal/vertical). Luminance was maintained constant in all the level of the stimuli (mean 146 RGB).

2. We have also amended the stimulus duration in the new version of the manuscript. As stimuli lasted for 5 frames, the presentation time was 59 ms.

3. Regarding the parameters for Bayesian t-test, we integrated that:

The analysis utilized the default prior settings: Prior Distribution for Effect Size: Cauchy, with a Scale Parameter of 0.707. We used these parameters for all the Bayesian t-test performed.

4. Regarding the statistical analysis employed to investigate the relationship between IAF and accuracy, we used a non-parametric multiple regression approach similar to that described in Cohen and Cavanagh (2011) and Samaha et al. (2017), which shared our very same aim of analyzing how trial-by-trail fluctuations of a neural marker is able to intercept trial-by-trial variability in decisional outcomes. However, we agree that complementing this analysis with generalized linear mixed effects models is a valuable addition. To this end, we fitted this type of model and confirmed the same pattern of results.
   In the new version of the methods, we added that:

   To corroborate the results, we examine the relationship between IAF and task accuracy by means of a generalized linear mixed model (GLMM). The dependent variable was task accuracy, coded as a binary variable (0 or 1). The normalized alpha frequency was used as a continuous predictor. To account for inter-individual variability, participants were included in the model as a random effect. The specific formula used for the model was:$accuracy \sim IAF + (1 \mid Subject)$*. MATLAB's fitglme function was used to fit the model to the data, specifying a binomial distribution and a logit link to reflect the binary nature of the accuracy data.*

   And in the supplementary material:

   GLMM analysis corroborated the IAF-sensitivity finding derived from the non-parametric multiple regression analysis (see main text) as IAF was a significant predictor of trial-by-trial accuracy (slope = 0.13, $p < 0.01$). This effect remained significant when accounting for trial-by-trial fluctuations in alpha power (slope IAF= 0.07, $p < 0.01$; slope alpha power = -0.03, $p = 0.14$).

5. We appreciate the opportunity to provide further clarification regarding the fitting procedures and parameter recovery of HDDM.
   Specifically, we included in the new version of the manuscript that:

   We used the HDDM package provided by (Wiecki et al., 2013)*. Here, Bayesian inference through Markov chain Monte Carlo (MCMC) sampling is utilized to approximate posterior distributions for model parameters at both the individual and group levels. This approach is advantageous as it effectively manages the uncertainty associated with parameter estimates, allowing for robust and precise parameter estimation. The priors for each parameter in HDDM are informed by a pool of 23 studies, which report the best-fitting parameters of the Drift Diffusion Model across various decision-making tasks* (Matzke and Wagenmakers, 2009)*. By grounding these priors in empirical data, HDDM ensures that initial parameter assumptions are reliable and informed by a substantial body of existing research (see the supplement by* (Wiecki et al., 2013) *for a visual representations of these priors).*

Reviewer #2 (Remarks to the Author):

This paper extends current proposal regarding perceptual sampling by the human alpha rhythm
by proposing that the relevance of alpha phase can be captured only at slower frequencies. This
is based on the idea that with faster rhythms, it is more likely that a stimulus will fall under a
'good' phase of the alpha rhythm. The mechanism is elegant and consistent with previous proposals. The methods are sound and overall, the paper is well written, albeit with some typos. I have the following comments.

We appreciate their overall positive feedback on our work.

Abstract and line 388: "Bayesian statistics, and advanced computational models" – I think that the authors are referring to using the drift diffusion model, and the word "advanced" is not necessary or useful (DDM was described by Ratfcliff and colleagues in 1978). Why not just write "drift diffusion model" which is the same number of words?
Abstract, I would tone down the "unprecedented" from the "we provide… insights into the neural mechanisms through which IAF influences perceptual decisions" – I think that the seeds of the idea have been in some of the writings by VanRullen, Busch and others, as well as in some papers from the 1930's onwards, and we expect most things published in this journal to be "unprecedented", so it is fine without the hype words.

Regarding these key points raised, we understand the concern relative to the term "advanced" in reference to the drift diffusion model, and the other regarding the perceived lack of novelty in linking IAF to perceptual accuracy.

1) We are open to removing the term "advanced". However, it's noteworthy that in the past 46 years (Ratcliff, 1978), no studies have integrated this data analysis

technique into this literature. Rather than advanced, we framed it in the new version of the manuscript as "previously unexplored".

2) We understand the reviewer's point that the relationship between IAF and perceptual sensitivity is not entirely unprecedented— we completely agree. However, no previous empirical research has investigated this phase-related mechanism linking IAF to perceptual sensitivity. Therefore, we strongly believe this represents unprecedented evidence. Nevertheless, we are open to removing this term if the reviewer and the editor agree it is appropriate to do so.

Abstract:

this sentence seems to be missing a word: "Higher probability for covering optimal alpha phases during same stimulus presentation, crucial for accurate perception, is intrinsic to faster than slower IAF." Perhaps "is intrinsic to faster rather than slower IAF"….

Thank you for pointing out the missing word. We have amended the abstract in accordance.

Methods:

Concerning the methods. It is established that alpha power, more consistently than frequency
or phase, has an effect on visual detection. Additionally, peak alpha frequency often covaries
with alpha power. Given the small magnitude of the effect reported it is imperative that the authors perform the same analysis they conducted by binning alpha frequency but by binning
alpha power. In other words, the author should clearly show that their results are not driven by
an effect in alpha power.

We fully agree with the reviewer on the importance of clarifying the specificity of the effects linking IAF to perceptual performance, ruling out the possibility that the same relationship could be driven by alpha power. To address this concern, we conducted a series of analyses that consistently demonstrated that alpha power does not significantly influence perceptual performance. Below is a summary of the analyses, which are detailed in the supplementary materials.

First, we extracted time-frequency amplitude maps and conducted a cluster-based permutation test to assess whether oscillatory amplitude differed based on response accuracy (correct vs. incorrect trials). The statistical analysis revealed no significant clusters distinguishing correct from incorrect responses ($p = 0.43$).
Next, we performed regression analyses using a least-squares approach, adding alpha power as an additional predictor alongside IAF. The results indicated that IAF remained a strong predictor of performance (slope = 0.37, SEM = 0.09; $t_{115} = 4.04$, $p < 0.001$, BF = 179.63), while alpha power (mean slope = -0.15, SEM = 0.09; $t_{115} = -1.57$, $p = 0.12$, BF = 0.34) did not demonstrate any predictive value. A similar pattern emerged when using a GLM model (mean slope for alpha power = -0.03, $p = 0.13$; mean slope for IAF = 0.08, $p < 0.001$).
We also conducted bin analyses by dividing trials into terciles based on alpha amplitude. The results showed no significant differences in sensitivity and criterion between trials in the first and third terciles (all $p > 0.09$, all BF < 0.41).
Finally, we conducted another time-frequency analysis comparing oscillatory amplitude between participants in the slow and fast IAF groups. No significant clusters were found between the two groups ($p = 0.14$).
In conclusion, these analyses consistently demonstrate that alpha power does not contribute to explaining perceptual performance, in contrast to the robust influence of IAF.

Here we report the new section added in the methods section:

Assessing the specificity of the relationship between IAF and Perceptual Accuracy relative to Alpha Power

To ensure that the effects observed in relation to IAF were not confounded by oscillatory power, we conducted several analyses (See Supplementary Materials). First, we extracted time-frequency oscillatory amplitude maps for each participant from the specific electrodes used to compute IAF. Time-frequency decompositions were performed for 50 linearly spaced frequencies ranging from 2 Hz to 50 Hz. Wavelets with 3 cycles at the lowest frequency and 11 cycles at the highest frequency were applied. After obtaining the absolute value of the resultant analytic signal, we conducted a cluster-based permutation test to evaluate whether oscillatory amplitude varied based on response accuracy (correct vs. incorrect trials). Next, we extracted the trial-by-trial amplitude from the time-frequency analysis, averaging it within the -800 ms to -100 ms window, at the frequency closest to each participant's previously determined IAF. We performed regression analyses using a least-squares approach, incorporating both trial-by-trial fluctuation in IAF and alpha power as predictors of perceptual accuracy. This analysis allowed us to determine whether the influence of IAF on performance was independent of alpha power. We also replicated a bin-based analysis by dividing trials into terciles based on the previously extracted trial-by-trial alpha amplitude. We compared signal detection theory metrics across the lowest and highest terciles to confirm that the effects observed with IAF did not hold when considering variations in alpha power. Lastly, we conducted a time-frequency analysis using the same parameters as those applied in the comparison of correct and incorrect responses. This analysis aimed to assess potential differences in oscillatory amplitude between participants categorized into low and high IAF groups.

Here we include the new part and figure included in the supplementary materials:

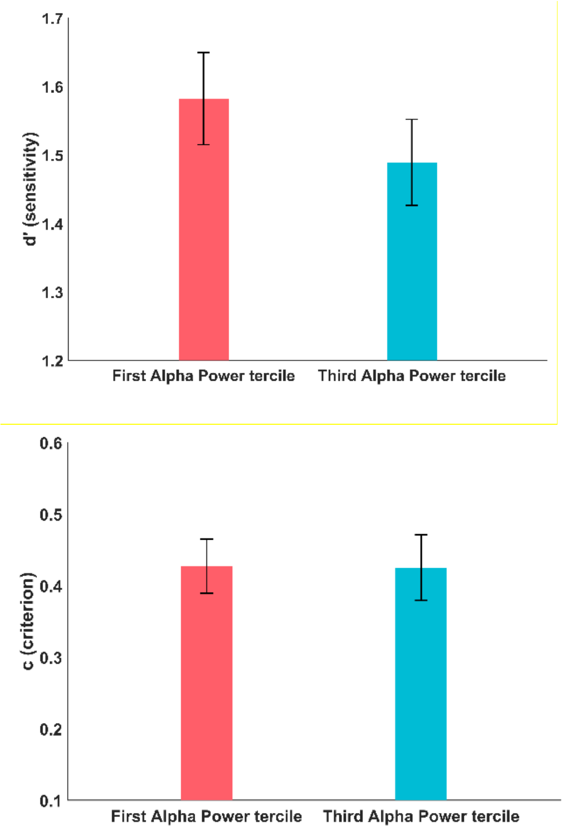


Figure S1. No significant differences observed in behavioral performance between first and third power terciles
We replicated the initial analysis included in the main text by binning trials based on pre-stimulus power in the first and third terciles. The analysis revealed that this binning approach did not dictate behavioural performance (all $t_{115} < 1.68$, all $p > 0.09$, all BF < 0.41).

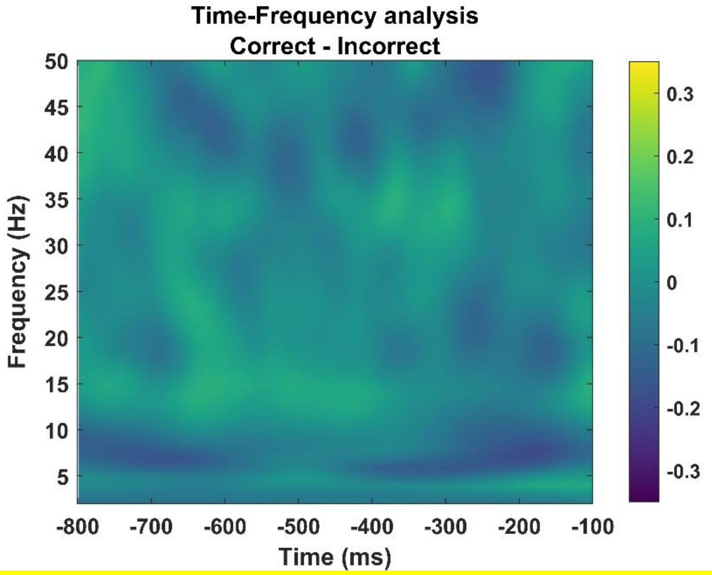


Figure S2. Correct and Incorrect trials are not characterized by difference in oscillatory amplitude.

We conducted a time-frequency analysis to assess whether correct responses exhibited different amplitudes compared to incorrect responses. The cluster-based analysis revealed no significant differences between the two types of trials, indicating that power, in contrast to IAF, is not capable of differentiating correctness in the trials.

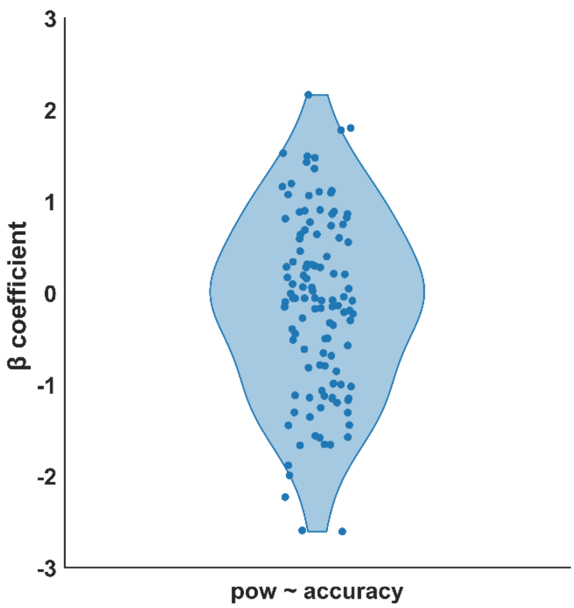


Figure S3. Trial-by-trial fluctuations in alpha power do not account for perceptual accuracy.

We conducted an analysis to evaluate whether trial-by-trial fluctuations in power, alongside those of IAF, influenced the accuracy of the responses provided. We extracted the power calculated from the time-frequency analysis by taking the alpha frequency closest to the IAF, normalizing it through a z-score transformation. Subsequently, we created a design matrix that included the intercept, IAF, and normalized alpha power. The regression analysis confirmed the significant predictive ability of IAF fluctuations on perceptual accuracy (mean slope = 0.37 $p < 0.001$). In contrast, fluctuations in power did not modulate the participant's ability to accurately retain the stimulus (mean slope = -0.15 $p = 0.12$).

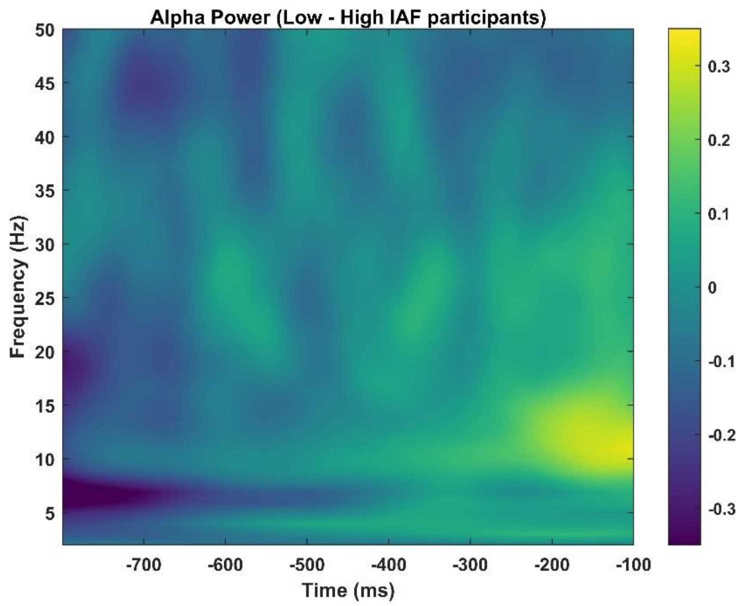


Figure S4. Low and Fast IAF groups do not exhibit significant differences in oscillatory amplitude.

We conducted a time-frequency analysis to assess whether participants in the low IAF frequency group exhibited different amplitude profiles compared to those in the high IAF frequency group. The cluster-based analysis revealed no significant differences between the two populations, indicating that power does not effectively differentiate between the two groups.

The manuscript seems to be a reanalysis of a previous dataset used in another paper published
by the same authors (https://doi.org/10.1162/jocn_a_02026). If this is the case, the authors should clearly acknowledge this through the manuscript and summarize the previous analyses and findings.

Thank you for your feedback. We appreciate the opportunity to clarify the relationship between the current manuscript and our previous publication (https://doi.org/10.1162/jocn_a_02026).

We would like to emphasize that the dataset analyzed in the present study is distinct from the one used in our previous publication. The earlier study focused on resting-state EEG data, whereas the current manuscript is based on task-related EEG data, which involves different experimental conditions and analytical approaches. Consequently, the analyses and findings presented here are completely independent of those in the prior publication.

However, for transparency and to provide context, we have acknowledged the relevant findings from our previous study that relate to the current research question. Specifically, we have referenced our prior observation that individuals with higher resting IAF require less contrast for accurate detection, as this finding aligns with the hypotheses tested in the present study. Furthermore, prompted by comment raised by this reviewer and reviewer 3, we conducted a similar analysis correlating mean pre-stimulus IAF with individual contrast thresholds. The results closely align with those from our analysis of resting-state IAF: individuals with higher IAF exhibited lower contrast thresholds (Pearson $r = -0.21$, $p = 0.03$; Pearson skipped $r = -0.21$, CI = [-0.37, -0.03]). We have incorporated this additional evidence linking IAF to perceptual sensitivity in the revised version of the paper. Specifically, in the introduction we acknowledged the previous finding:

> *Additionally, we have recently demonstrated that individuals with a higher resting IAF require less contrast for accurate detection compared to low IAF participants* (Tarasi and Romei, 2023).

While in the results we integrated the new finding:

> *Building on our recent finding that individuals with higher resting IAF require less contrast for accurate detection compared to those with lower IAF*(Tarasi and Romei, 2023)*, we further demonstrate that this relationship extends to average pre-stimulus alpha levels. Specifically, individuals with higher pre-stimulus IAF consistently*

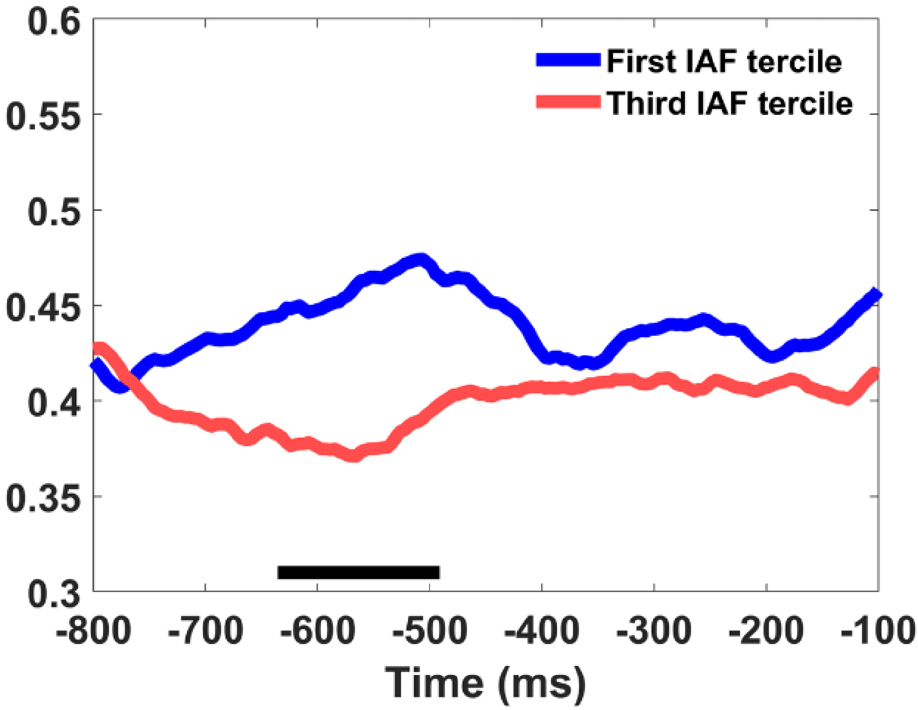
required lower threshold contrast to achieve target performance levels (Pearson = -0.21, p = 0.03; Pearson skipped = -0.21, CI = [-0.37, -0.03]).

Discussion:

Results

Criterion – perhaps there is some weak effect at specific time points that is missed by correcting over the whole time period? Other studies have reported criterion changes with alpha power, so it would be good to be careful about how this is worded.

In the initial version of the manuscript, we noted the absence of a significant cluster (i.e., contiguous time points) that survived correction for multiple comparisons. However, as per reviewer observation, there were specific time points in which a difference in criterion between the two types of trials emerged. Notably, the largest cluster, while statistically not significant, had a p-value of 0.08. However, upon applying a more refined filter and utilizing a traditional correction for the SDT-related indices in order to comply to the request of reviewer 1, we identified a specific as well as narrow time period (from -640 to 500ms) where variations in IAF significantly influenced the criterion.
This result suggests that in trials with higher IAF, there appears to be a reduction in response bias. In the Signal Detection Theory (SDT) framework, a criterion of 0 signifies no bias, positive values reflect a conservative bias, and negative values indicate a liberal bias. As illustrated in Figure 1C (attached below), participants' criteria tend to shift toward positive values, suggesting a conservative bias in their responses. This finding is consistent with the well-established observation that individuals often exhibit a preference for conservative over liberal criteria (Rahnev and Denison, 2018). However, in the highlighted cluster within the figure, it becomes clear that during trials with high IAF, the criterion value decreases (i.e., approached 0), suggesting that participants display less bias compared to trials with lower IAF.



Still, the observed effect linking IAF and criterion likely results from an indirect influence exerted by IAF. Specifically, decision bias decreases in trials with higher IAF, which are also trials where perceptual sensitivity is high. This criterion shift could, therefore, stem from an enhanced ability to perceive the target, making participants less reliant on an individual tendency to report (or not report) the target. This assertion is further supported by the negative relationship we identified between task sensitivity and criterion: as participants' accuracy increases, their response bias decreases (Pearson = -0.30, $p < 0.01$; Pearson skipped = -0.24, CI = [-0.39 ; -0.06]). To further demonstrate that the relationship between IAF and response bias is likely spurious (already clear in our opinion, as this is the only analysis in which a small IAF-criterion relationship emerged), we conducted three new analyses that clearly demonstrate there is no actual relationship between IAF and bias:

1. We conducted an analysis paralleling the one that demonstrated differences in IAF values between correct and incorrect trials, with correct trials displaying higher IAF. In the new analysis, we grouped trials based on choice (i.e., report stimulus presence vs. stimulus absent) rather than response accuracy (i.e., correct vs. incorrect choice). The rationale here is that, if a higher IAF were driving a tendency

to report the presence versus absence of the target (i.e., indicating a perceptual bias), we would expect to find a significant relationship between trial type and IAF. However, IAF did not vary with choice; it remained constant regardless of whether the target's presence or absence was reported.

2. Similarly, we repeated the trial-by-trial analyses, entering choice (rather than accuracy) as the dependent variable, to investigate whether IAF fluctuations could predict choice variations. Once again, IAF fluctuations did not predict whether participants reported the target as present or absent.
3. Lastly, we developed a new trial-by-trial analysis including both accuracy and choice as predictors, using IAF as the dependent variable. The results showed that only accuracy predicted IAF (with correct trials associated with higher IAF), while choice—reflecting any potential response bias to report target presence versus absence—did not predict IAF.

With these three additional analyses (alongside those from the original submission), we now have comprehensive evidence that IAF and bias are unrelated. We report these additional analyses in the supplementary materials:

To further demonstrate that the relationship between IAF and response bias is likely spurious (as indicated by the fact that a small IAF-criterion relationship emerged in only one analysis), we conducted three additional analyses that collectively support the absence of a genuine IAF-bias relationship:

1. *We repeated the analysis included in the main text that showed higher IAF values in correct compared to incorrect trials. In this new analysis, however, trials were grouped by choice rather than accuracy. The rationale was that if higher IAF promoted a tendency to report target presence over absence (i.e., reflecting a perceptual bias), a significant relationship should emerge between choice type and IAF. However, IAF did not vary with choice; it remained stable whether participants reported the target's presence or absence (IAF report present = 11.35 ± 0.09, IAF report absence = 11.33 ± 0.09, t_{115} = 1.60, p = 0.11, BF = 0.34).*

2. *We also repeated the trial-by-trial analysis using choice (rather than accuracy) as the dependent variable, examining whether IAF fluctuations predicted variations in choice. Again, IAF fluctuations did not predict whether participants reported the target as present or absent (mean slope = 0.12, p = 0.21, BF = 0.22).*
3. *Lastly, we performed a new trial-by-trial analysis including both accuracy and choice as predictors, with IAF as the dependent variable. This analysis showed that only accuracy predicted IAF (mean slope = 0.41, p < 0.001, BF = 262.49), with correct trials associated with higher IAF, while choice—indicating any potential bias towards reporting target presence—did not predict IAF (mean slope = -0.06, p = 0.52, BF = 0.12).*

Figure 4C – It is interesting to note that the significant bins for ITPC in slow IAF participants, which are not found for the fast IAF individuals, is in the high alpha to low beta range (appears to be 10-18 Hz approximately from the figure, at least). It would be good to point this out, rather than simply describing it as "confined to the alpha band" (line 294), which does not seem quite accurate.

Thank you for your observation. We have revised the main text to more accurately reflect the results, noting that the significant bins for ITPC in slow IAF participants are in the high alpha to low beta range, rather than just describing them as "confined to the alpha band." This adjustment ensures the text is consistent with the findings shown in the figure.

I agree that drawing a clear connection between studies showing a role of alpha in detection (cited in the paper) and studies showing a role of alpha in temporal integration is a valuable theoretical contribution. In addition, the ideas put forward could be tested, and I think that it would be useful to make these implications clear with a few examples. For example, the role of alpha phase and IAF on detection should vary depending on stimulus duration (see Figure 5). Second, increasing IAF externally via entrainment/neurostimulation could improve detection while also decreasing the role of phase. Given that IAF varies naturally between people and as a function of age and neural disorders (such as SCZ as mentioned by the authors),

Thank you for the comment. We agree that highlighting the implications and testable hypotheses is a valuable way to clarify our proposal and stimulate further research in this field. Below is our response to your thought-provoking suggestions:

We have conducted a deep consideration of the relationship between stimulus duration and its potential impact on our findings. It is reasonable to posit that, in a particular context like the one employed in our experimental protocol, shorter stimulus durations, particularly those significantly less than one alpha cycle, accentuate the role of phase angles in determining response accuracy. Specifically, when stimuli are presented for brief intervals, the timing of their onset relative to the ongoing alpha oscillations can significantly influence perceptual outcomes. In such cases, the precise timing of stimulus presentation in relation to the alpha phase becomes critical for perceptual performance. Higher intrinsic IAF enables the brain to rapidly sample a broader range of phase angles. This enhanced sampling capability increases the likelihood of encountering advantageous phases, while lower IAFs restrict the diversity of sampled angles, thereby diminishing the probability of achieving the optimal angles necessary for accurate perception. Thus, a slower IAF may lead to heightened phase dependency, as fewer angles are available within the same temporal frame.

However, in light of the reviewer's insightful comments, we recognize the necessity of introducing another pivotal parameter into our model: the variability of sensory input. We propose that the relationship between IAF and phase is strongly moderated by the variability of the stimulus. Presenting a static stimulus for an extended duration can mitigate the influence of phase on perceptual outcomes. However, while this consideration is particularly pertinent in controlled experimental settings, it does not fully incorporate the complexities of real-world experiences, where stimuli are dynamic and continuously changing. In natural environments, perception is constantly modulated by various factors, including variations in sensory inputs, spatial shift, eye movements, saccadic fluctuations, as well as environmental conditions such as lighting and background noise. The interplay of these factors suggests that real-world perception could not rely just on the integration of the immediate properties of stimuli but also on the broader context and temporal properties in which they are experienced.

Referring to discrete perception theory, which posits that our representation of the external world is constructed from a series of snapshots sampled by the cognitive system (Van Rullen, 2016), we can better understand how the IAF-phase relationship may assume a central role in these scenarios. The quality of a coherent and accurate representation of the external world, based on the integration of these brief snapshots, is directly contingent on sampling speed (Samaha et al., 2015). Thus, the phase during which these "short-lasting" snapshots are taken is crucial for determining their fidelity. If a snapshot is captured during a favourable phase, its accuracy is markedly enhanced compared to a snapshot taken when the phase is unfavourable. This implies that the timing of sensory sampling in relation to ongoing neural oscillations can significantly influence the fidelity of perceptual representations. Hence, the effects of IAF and phase may accumulate in a dynamic reality, where the ability of a higher IAF to favour entry into advantageous phases aggregates as various snapshots are integrated over time. This cumulative effect may contribute to a more accurate and coherent representation of the external world.

In this context, analogy with digital samplers could be valuable. Consider two samplers operating at different frequencies: one at 70 Hz and the other at 130 Hz. If a phenomenon is relatively static, both samplers may yield similar representations of the generating source. However, in scenarios involving rapid changes (as the sensory environment in which we are embedded), the 70 Hz sampler may miss critical fluctuations that the 130 Hz sampler can capture. Therefore, a slower IAF limits the brain's capacity to sample rapid and subtle changes in stimuli, leading to less accurate or even distorted representations of external reality. Moreover, the 70 Hz sampler will be more reliant on the precise timing of when a phenomenon occurs; under particularly favourable conditions, its resulting signal representation may closely resemble that produced by the 130 Hz sampler. Nonetheless, the latter increases the probability of detecting minimal changes inherent in the external signal.

Another crucial aspect to consider is how perceptual decision-making unfolds. According to the influential accumulator-to-bound theory (Gold and Shadlen 2007), the perceptual system samples incoming stimuli until a decision threshold is reached. This suggests that perception doesn't rely on a single sampling of the stimulus to determine the outcome;

instead, it involves a dynamic process of multiple samples until a decision is made. Consequently, even when the stimulus is static, a faster sampler has a distinct advantage in capturing external information more efficiently. This efficiency allows it to reach the decision threshold more reliably, less affected by noise. Indeed, increased sampling enhances the capacity to extract the true signal from noise, whereas fewer samples make it more likely that incidental noise during accumulation might drive the decision process towards an incorrect threshold. Therefore, even in cases where a static stimulus lasts for an extended period, the decision-making process remains more efficient with a faster sampler.

Furthermore, in this context, the example of schizophrenia (SCZ) could be of relevance as well. Individuals with schizophrenia exhibit significant perceptual disturbances and difficulties in accurately representing external reality (Butler et al., biological psychiatry 2008; Fan et al., schizophrenia 2024), with seminal studies indicate that the incidence of disruption in visual perception among schizophrenia patients is high, ranging from 40% to 62% (Phillipson, psychological medicine 1985). This symptomatology could depend on the reported slowing of alpha oscillations in SCZ which may limit the ability to effectively sample the environment (Murphy and Öngür, 2019; Trajkovic et al., 2021; Tarasi et al., 2022). This phenomenon is highlighted by Ramsay et al. (2021), who demonstrates that the slow down in IAF among schizophrenia patients translates into a greater likelihood of experiencing distorted or inaccurate perceptions. Thus, a pathological slowing of IAF may not only compromise immediate perception but would also scale down to the overall ability to integrate information from the surrounding environment. Returning to our hypothesis, this limitation could reduce the probability that sensory information will be sampled during favourable phases, leading to a global representation of reality that is disrupted and misaligned as observed in schizophrenia spectrum. While this reasoning could be seen as a long leap, we echo the request of the reviewer to provide context for testable hypotheses. Within this context, individuals with schizotypy or schizophrenia may represent a relevant test population in which to ascertain this hypothesis.

In conclusion, we state that the interaction between IAF and phase is fundamental in determining the quality of perceptual outcomes. A higher IAF enhances the ability to navigate through a broader range of phase angles, facilitating improved performance, whereas a slower IAF impedes this process, resulting in suboptimal sampling of the external environment. The dynamics and duration of the external stimulus play a moderating role in this relationship: static and prolonged stimuli diminish the effect of phase, while dynamic stimuli are associated with an additive phenomenon, an hypothesis we are currently testing empirically in our lab. In this context, the series of snapshots acquired from the external environment would more likely be influenced by the phase at which the stimulus enters the perceptual system. A lower IAF is linked to increasingly poor representation at a probabilistic level due to the reduced likelihood that each sampling captures the stimulus (and its dynamic changes) during favourable phases. Conversely, higher IAFs can encompass a greater range of phase angles, resulting in an accumulation of snapshots that have been appropriately processed, ultimately leading to a faithful representation of the external signal.

We have incorporated this reasoning into the introduction, stating that:

> *The basic idea we propose here is that the mechanism facilitating a more accurate response with faster (relative to slower) IAF lies in its capacity to span a broader distribution of phase angles within the same stimulus presentation timeframe. As a consequence, this heightened phase coverage increases the likelihood of aligning with optimal phase angles, crucial for the accurate perception of stimuli. This proposition, in turn, leads to another hypothesis closely intertwined: the ability of phase angles in dictating the accuracy of the response is expected to be more pronounced in participants or trials in which IAF is slower. In such instances, the reduced likelihood of falling into optimal phases during stimulus processing makes the phase angles a crucial determinant of performance (i.e., if the phase is in non-optimal angles, the likelihood of reaching the optimal ones will be reduced). On the contrary, when alpha is faster, it aids in covering numerous phase angles during stimulus presentation, thereby increasing the likelihood of processing stimuli within those optimal phases.* ***This relationship is particularly relevant given the presentation time of our stimuli (i.e., 60 ms), where the phase position becomes more critical due to the limited time available to cover a wide distribution of phase angles.***

Moreover, in the discussion we added that:

> Another important aspect to consider and further investigate is the role of stimulus duration, as it could significantly influence the interaction between IAF and alpha phase. For brief stimuli—specifically those lasting less than one alpha cycle, such as those used in our study—the interaction between IAF and phase may be more pronounced compared to longer stimuli. However, another crucial factor modulating this relationship is the variability of sensory input. Since the external world is inherently dynamic rather than static, the effects of IAF and phase may accumulate across this fluctuating environment. In such context, higher IAF would make it more likely to enter favourable phases, with these effects aggregating as subsequent snapshots of sensory information are integrated over time. This cumulative process could contribute to a more accurate and coherent representation of the external world. Furthermore, it is also conceivable that, even for long-lasting static stimuli, this mechanism could exert a similar influence. In these cases, the increased opportunity to accumulate more information within a given time period could enhance the statistical reliability of sensory processing, with more frequent updates facilitating a robust integration over time. Follow-up studies incorporating a range of stimulus durations will be essential to determine whether the observed phase effects are consistent across different temporal contexts, or if they are specific to the short-duration stimuli used in our study. Moreover, varying stimuli over time may offer a more complex dynamic of the interplay between IAF and phase angles, a hypothesis currently tested in our laboratory.

Relating to the second point you raised, the evidence presented in our study is correlational, reflecting the nature of our protocol. We do agree that follow-up studies should investigate whether modulating IAF through neurostimulation or sensory entrainment protocols could alter the influence of phase on perceptual performance. Based on our findings, we hypothesize that increasing IAF might be associated with a reduction in the impact of phase on performance, whereas a decrease in IAF could amplify the role of phase in influencing perceptual accuracy.

We integrated this hypothesis into the discussion:

> *Building on this reasoning, we hypothesize that altering IAF through neurostimulation* (Di Gregorio et al., 2022; Bertaccini et al., 2023; Trajkovic et al., 2024) *or sensory entrainment protocols* (Ronconi et al., 2018) *could further validate this interaction. Specifically, increasing IAF is expected to reduce the impact of alpha phase on perceptual accuracy, while decreasing IAF might enhance the role of phase in shaping perceptual performance.*

Regarding the last point raised by the reviewer, we believe (the phrase is truncated) the reviewer was referring to the variability in IAF among individuals and its influence on perceptual performance, particularly considering factors such as age and psychopathological conditions. Variations in IAF across individuals present a valuable opportunity to explore how these differences impact perceptual performance and how the relationship between IAF and alpha phase might vary. In the revised manuscript, we have expanded on how these inter-individual differences contribute to understanding the dynamics between IAF, alpha phase, and perceptual outcomes.
As highlighted by the reviewer, age-related decline in IAF has been consistently reported (Klimesch, 1999), which would aid in exploring how the IAF-phase relationship varies across different age groups. Genetic predispositions are another factor influencing IAF (Smith et al., 2006).
Moreover, psychopathological conditions, such as schizophrenia (Ramsay et al., 2021), disrupt typical IAF dynamics and then could alter IAF-phase interactions. Additionally, white matter structure plays a crucial role in IAF variability. For instance, the integrity of white matter tracts connecting the thalamus and primary visual cortex modulates IAF (Minami et al., 2020). This structural variability could impact how phase interacts with IAF, potentially accounting for some observed differences in our study.

We have incorporated these insights into the discussion of the revised manuscript to emphasize the significance of IAF variability stemming from various factors and its implications for investigating the mechanisms we have identified.

> *Our findings underscore the importance of considering individual differences in IAF when interpreting phase-related effects. Inter-individual IAF variability is influenced by several factors, which could, in turn, modulate the IAF-phase relationship found. For instance, the properties of the optic radiation—the primary white matter tract connecting the thalamus to the primary visual cortex—as well as genetic factors play a role in modulating individual differences in IAF (Minami et al., 2020; Smith et al., 2006). IAF also tends to decline with age (Klimesch, 1999), and conditions such as schizophrenia are associated with a lower IAF (Ramsay et al., 2021). Future research should explore how these IAF-related factors might modulate the role of phase in perceptual processes, either enhancing or diminishing its impact.*

It would be good to be more precise with the use of the word "precision". For example: "In alignment with these findings, our results strongly contribute to this body of research as they indicated a significant role of IAF in dictating the precision of sensory acquisition, as spontaneous inter-trials fluctuations in IAF were able to strongly predict the accuracy of perceptual decision." As the authors note, they only measured accuracy and not precision of sensory representations since it was a simple detection (present/absent) task. It may indeed be the case that IAF might influence the amount of time that a brief stimulus is sampled during optimal phase, but the authors did not show that this is the case or that this decreased time necessarily reduces the precision of the neural response or perceptual interpretation. Again, this could be tested with a different paradigm, leading to a strong hypothesis that a task in which duration of sensory sampling improves precision is influenced by IAF and alpha phase.

Thank you for this valuable observation. We agree that, given the present paradigm, it is indeed not feasible to directly measure how IAF influences the precision of sensory representation. In light of this, we have adopted more descriptive terminology to better capture our findings. Specifically, we now refer to the modulation of *perceptual sensitivity* by IAF and its interaction with phase. For example, we changed the part of the text highlighted stating that:

> *In alignment with these findings, our results strongly contribute to this body of research as they indicated a significant role of IAF in dictating* ***perceptual sensitivity****, as spontaneous inter-trials fluctuations in IAF were able to strongly predict the accuracy of perceptual decision.*

The authors suggest that their model could explain null findings based on the fact that previous
research has overlooked the importance of peak alpha frequency when looking at the effect of
phase. Similar proposals have been advanced. For example, it has been shown that other intervening variables could explain the effect (or lack thereof) of alpha phase/frequency. Particularly, consideration of alpha power (Fakche et al. 2022 eNeuro) and/or aperiodic properties of the EEG (Deodato & Melcher, 2024, biorkiv) has been shown to affect result. These alternative viewpoints could enrich the discussion.

Thank you for pointing out these recent works, which are very much in line with the topic of the work. We have integrated them into the main text:

> *Additionally, this evidence aligns with studies suggesting that the explanatory power of brain rhythms on behavior is enhanced when combining multiple oscillatory indices, rather than examining them separately. For instance, Fakle et al.* (2022) *provide compelling causal evidence demonstrating that the effect of alpha phase on TMS-induced phosphenes is more pronounced during high alpha amplitude trials, while it is less influential during low alpha amplitude trials. Additionally, integrating aperiodic components could offer further insights. For example, Deodato and Melcher* (2023) *have shown that, alongside alpha frequency, the aperiodic component also impacts performance in the flash fusion task.*

We believe that the proposal is now even more compelling, as, in addition to the IAF-phase relationship, other studies have demonstrated that combining multiple oscillatory indices explains behavior more effectively than considering them in isolation.

General: please use spellcheck for minor errors such as "temporal integratrion" and so on. I think that simply using Word to check the text would help.

We have corrected any typographical errors. Thank you for the note.

Other comments:
Lines 56-57 if the authors were the first to postulate such hypothesis, they should add an appropriate reference. Otherwise appropriate references could be:
Valera et al. 1981
Kristofferson 1967

Thank you for suggesting these references that we integrated in the main text.

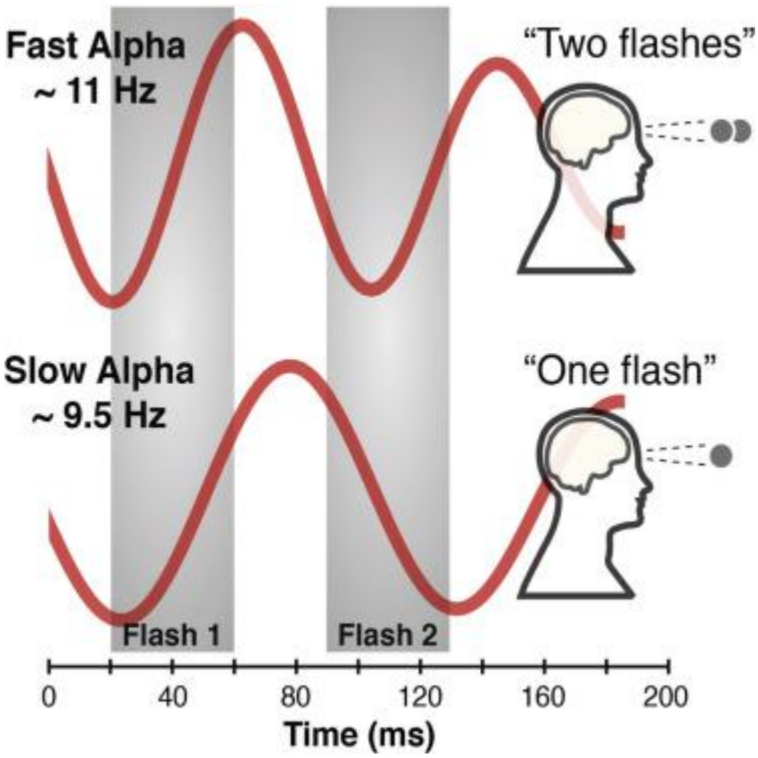
Lines 62-69 present an oversimplified view of the effects of neurostimulation on temporal integration windows. Many researches using entrainment or neurostimulation reported effects that are not entirely supporting this view. This should be acknowledged.

Thank you for your comment. To offer a more comprehensive perspective, we have included additional references (Gray and Emmanouil, 2020; Ronconi et al., 2022; Mokhtarinejad et al., 2024) in the revised manuscript that address studies reporting null or mixed results regarding the impact of neurostimulation on temporal integration. We think that these inclusions provide a more balanced overview of the current research.

Lines 83-87 "This leads to the intriguing hypothesis that variations in alpha frequency play a more basic role in visual processing beyond its temporal aspects, influencing visual sensitivity and extending even to low-level stimulus features."
It's not clear how one thing would lead to another, the authors need to make a better point at this. This seems to be one of the key points in the paper. However, it is presented in an

approximate way.

Thank you for your observation. We realize that our initial explanation was not entirely clear. Our central hypothesis is that the relationship observed in our study, where IAF correlates with greater accuracy in stimulus detection, is underpinned by the same mechanism that explains the link between IAF and temporal resolution. In the well-known 2-flash fusion task, it has been suggested that stimuli falling within the same alpha cycle are integrated, while those falling in different cycles are segregated. This can be explained by considering the alpha cycle as a basic sampling unit for external stimuli. If two flashes occur within the same alpha cycle, they are sampled together and thus perceived as a single event. In contrast, flashes that fall in different cycles are processed separately, leading to their perceptual segregation (Samaha and Postle, 2015). The proposal is the same mechanism, which explains the relationship between IAF and temporal resolution, also governs the relationship between IAF and perceptual resolution as identified in our study. Similar to the two flashes in the fusion task, if a perceptual stimulus (e.g., a checkerboard pattern) spans two separate alpha cycles, it will be sampled twice. If it falls within a single cycle, the perception of the stimulus decreases because the input is processed only once. Additionally, the influence of alpha phase and the IAF-phase relationship extends beyond perceptual resolution to also affect temporal resolution. A higher alpha frequency increases the probability that two flashes will fall within the good phase of two alpha cycle, whereas lower frequencies make this less likely. Below, we have attached the graphical abstract from Samaha and Postle (2015). If we hypothesize that in this example the optimal alpha cycle phase for processing is from the alpha peak to the alpha peak - half cycle (but same logic holds true if considering other configurations), we can deduce that a lower IAF only aligns with this optimal phase for the first flash of a stimulus. In contrast, a higher IAF is able to sample the stimulus during the optimal phase across two cycles. This difference would result in an integration response with higher IAF and a segregation response with lower IAF.



Therefore, our hypothesis is that the same underlying mechanism explains both perceptual accuracy in detection tasks and temporal resolution in flash fusion task.

To clarify this concept, we have revised the introduction to include the following explanation:

> *Here, we posit that alpha rhythm may serve as a general neural code underpinning perceptual sampling, influencing both temporal and perceptual resolution. The underlying mechanism can be understood through the classic 2-flash fusion task. In this task, authors have proposed that when two flashes of light are presented within the same alpha cycle, they are perceived as a single flash because they are integrated within the same temporal window. Conversely, when the flashes fall in different alpha cycles, they are perceived as two distinct events, due to separate sampling moments* (Samaha and Postle, 2015). *Building on this concept, we hypothesize that the same mechanism governs the relationship between IAF and perceptual resolution. For example, if a visual stimulus is presented in a way that it spans across two different alpha cycles, it will be sampled twice by the brain,*

> *leading to a clearer and more distinct perception. However, if the stimulus falls within a single alpha cycle, it is processed only once, potentially reducing perceptual performance.*

Lines 97-99. "This proposition, in turn, leads to another hypothesis closely intertwined: the ability of phase angles in dictating the accuracy of the response is expected to be more pronounced in participants or trials in which IAF is slower."
This is true only for short lived stimuli. Markedly shorter than the length of an alpha cycle but
not too short. The authors should specify in introduction and discussion the importance of stimulus length on the proposed mechanism.

As we have discussed in detail in a previous comment, we fully agree that stimulus duration is an important element to incorporate into our model. This factor plays a crucial role in the IAF-phase relationship, though it may be in turn modulated by the variability of external stimulation. We hypothesize that when a stimulus is dynamic, as is typically the case with external inputs), its accurate analysis inherently depends on detecting its subtle changes evolving over time. If the sampler that acquires snapshots from the external environment is more precise—such as when IAF is higher—there are more opportunities to capture the stimulus evolution during favourable phases, potentially enhancing perceptual sensitivity even for longer, yet variable stimuli. We are currently investigating this relationship in our lab, assessing whether the variability of stimulation is indeed a critical factor in modulating the impact of different frequencies within the alpha band on performance accuracy.
We have incorporated these considerations into the manuscript and thank the reviewer again for this valuable suggestion.
In the introduction we added:

> *The basic idea we propose here is that the mechanism facilitating a more accurate response with faster (relative to slower) IAF lies in its capacity to span a broader distributions of phase angles within the same stimulus presentation timeframe. As a consequence, this heightened phase coverage increases the likelihood of aligning*

> *with optimal phase angles, crucial for the accurate perception of stimuli. This proposition, in turn, leads to another hypothesis closely intertwined: the ability of phase angles in dictating the accuracy of the response is expected to be more pronounced in participants or trials in which IAF is slower. In such instances, the reduced likelihood of falling into optimal phases during stimulus processing makes the phase angles a crucial determinant of performance (i.e., if the phase is in non-optimal angles, the likelihood of reaching the optimal ones will be reduced). On the contrary, when alpha is faster, it aids in covering numerous phase angles during stimulus presentation, thereby increasing the likelihood of processing stimuli within those optimal phases.* ***This relationship is particularly relevant given the presentation time of our stimuli (i.e., 60 ms), where the phase position becomes more critical due to the limited time available to cover a wide distribution of phase angles.***

Moreover, in the discussion we added that:

> *Another important aspect to consider and further investigate is the role of stimulus duration, as it could significantly influence the interaction between IAF and alpha phase. For brief stimuli—specifically those lasting less than one alpha cycle, such as those used in our study—the interaction between IAF and phase may be more pronounced compared to longer stimuli. However, another crucial factor modulating this relationship is the variability of sensory input. Since the external world is inherently dynamic rather than static, the effects of IAF and phase may accumulate across this fluctuating environment. In such context, higher IAF would make it more likely to enter favourable phases, with these effects aggregating as subsequent snapshots of sensory information are integrated over time. This cumulative process could contribute to a more accurate and coherent representation of the external world. Furthermore, it is also conceivable that, even for long-lasting static stimuli, this mechanism could exert a similar influence. In these cases, the increased opportunity to accumulate more information within a given time period could enhance the statistical reliability of sensory processing, with more frequent updates facilitating a robust integration over time. Follow-up studies incorporating a range of*

stimulus durations will be essential to determine whether the observed phase effects are consistent across different temporal contexts, or if they are specific to the short-duration stimuli used in our study. Moreover, varying stimuli over time may offer a more complex dynamic of the interplay between IAF and phase angles, a hypothesis currently tested in our laboratory.

Line 145 "perceptual decision-making perception" it's likely a typo.

Thank you for the note. We amended the typo.

Lines 173-174. There seems to be an inconsistency between what the authors reported (-800 to
-450) and the figure 1.

Thank you for pointing this out. The issue was with the figure, which showed incorrect timepoints. We have now ensured consistency between the text and the figure (now this image is included in the figure 2).

Lines 193-194, The effect is very small (it depends on the decimal places), this is concerning.

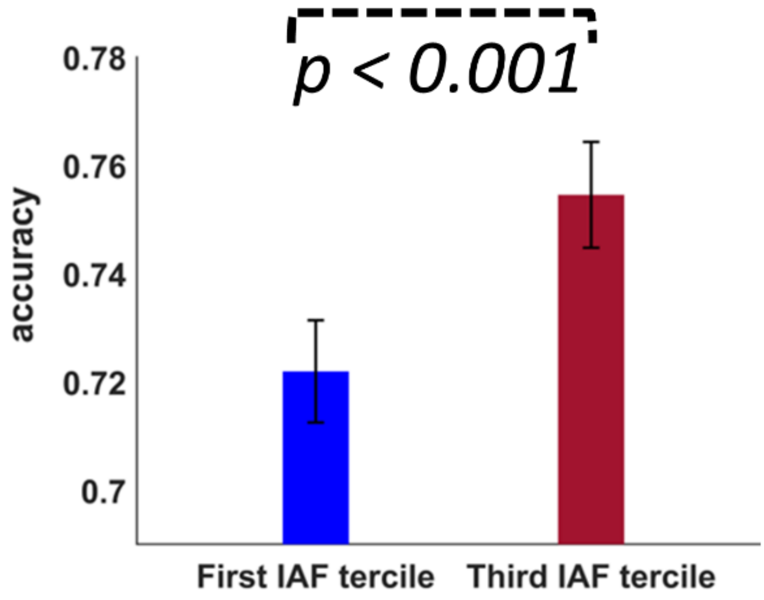
We thank Reviewer 2 for the comment. We also noted that Reviewer 1 raised a similar issue, highlighting how this difference is numerically small. However, we have a set of arguments organized into five distinct points that, in our view, convincingly support the validity of the observed effect. Here we replicate part of the response given to Reviewer 1:

We think that it is important to clarify that the mean observer represents an arithmetic aggregate of mean values both within and across participants. Indeed, the variability reported by looking at the "mean observer" does not represent the range of variability within and between participants, and as such it hampers the impact that instead bin-based and trial-by-trial analysis provided. The mean observer effect was reported not as a main finding but as a confirmatory finding supported by strong frequentist and Bayesian effect statistics. It just means that intraindividual and interindividual variability are important mediators of this effect as shown in the set of analyses provided. Indeed, despite the intra- and inter- individual variability, the small difference noted in this analysis between correct and incorrect responses does not diminish the findings. On the contrary, this result is exactly what one might expect given the nature of the phenomenon, it is therefore highly plausible and reflects previous effect sizes by seminal works (see below). Below we present 5 key points to support this statement:

1) **Within-Individual Variability**: The observed difference in IAF between correct and incorrect trials might be small at the arithmetical level because the phase coverage is quite similar within the same individual. To illustrate this, we calculated the standard deviation of IAF across time, trials, and participants, resulting in a value of 0.89 Hz (strikingly close to the value of 0.9Hz highlighted by the seminal work of Haegens et al., NeuroImage, 2014). Therefore, for a participant with an IAF of 10 Hz, the range of IAF values will fall between 8.2 and 11.8 Hz in 95% of cases (MEAN ± 2 STD). Comparing these two extreme cases, the stimulus coverage would be 49% for an IAF of 8.2 Hz and 71% for an IAF of 11.8 Hz. This important difference in coverage is in line with the hypothesis that a higher IAF significantly increases the likelihood of accurate responses compared to a lower IAF due to this differential phase coverage. However, in other scenarios where the within-subject IAF difference is less pronounced between different trials, the coverage would show less variation, as the reviewer correctly pointed out. For example, with a 0.2 standard deviation difference (10 Hz - 0.2 * 0.89 = 9.82 Hz; 10 Hz + 0.2 * 0.89 = 10.18 Hz), the coverage would be 58.9% for an IAF of 9.82 Hz and 61.1% for an IAF of 10.18 Hz. Therefore, in this case, the likelihood of the stimulus falling within the optimal phase is similar.
   Connected to this point, it has to be considered the strong impact of intra-participant averaging procedure we employed. When averaging IAF within each participant across all time points (800 points) and trials (150 trials), this process smooths out inter-time and inter-trials fluctuations, reducing the variability of IAF values. Differences that might be significant at the single trial and timepoint become less pronounced when averaged across numerous points and trials. This step alone

reduces the impact of any extreme IAF values by bringing them closer to the participant's mean IAF.

Crucially, we have evidence validating this hypothesized scenario empirically. In fact, recognizing the significant impact of the averaging procedure, we incorporated a bin-based analysis in the initial version of the manuscript (see image attached below). This analysis divided trials into the 1st and 3rd bins based on individual alpha frequency (IAF) values on a subject-by-subject basis. Crucially, this analysis reduces the smoothing effect of averaging at the intra-individual level, allowing us to consider the trials in which the deviation from the mean was consistent (on average, above (3^{rd} bin) vs. below (1^{st} bin) 1 SD from the mean). This is proved by the simple observation that in the first bin, the IAF average across participants was 10.90 Hz, while in the third bin, it was 12.19 Hz. So, by smoothing out the effect of the mean, the difference is around 1.3 Hz between the first and the third terciles. This value is about 26 times larger than the difference of 0.05 Hz. In this case, on average, in the first bin the alpha cycle covers approximately 65.4% of the stimulus, while in the 3rd bin the alpha cycle covers approximately 73.2% of the stimulus. Considering 116 individuals with 150 trials each, the difference created a notable impact. Indeed, the difference in phase coverage is compatible with the highlighted behavioral effect, in which the average accuracy in the third bin is 3 points higher than the accuracy measured in the first bin.



2) **Moreover, another crucial step to consider is the impact of inter-participant averaging**. After averaging IAF values within each participant, these averaged values have been further averaged across all 116 participants. This inter-participant averaging further reduces variability. For all these reasons, differences in IAF between correct and incorrect responses that were apparent within individuals are less likely to stand out in the grand average. The averaging process integrates a wide range of individual IAF differences, encompassing varying degrees of response accuracy and inherent IAF variability. Consequently, the differences between correct and incorrect responses observed at the individual level tend to converge towards the mean when averaged over a large number of participants. **Indeed, this process exemplifies regression toward the mean**, a statistical phenomenon where extreme values tend to move closer to the average upon repeated measurements or averaging. By averaging IAF values within and across participants, extreme differences are moderated. As a result, the final estimate shows smaller differences in IAF between correct and incorrect responses than might be seen in individual trials. Consequently, the overall IAF differences are less pronounced in this type of analysis, as the repeated averaging process obscures the more extreme values and highlights those closer to the mean IAF. This is precisely the reason that led us to also adopt trial-by-trial analysis that are able to

weigh the single-trials impact, avoiding averaging procedure (indeed, this procedure allowed us to find an even greater Bayes Factor in the trial-by-trial variability that was however already high in the mean value analysis).

Moreover, in acknowledging the influence of inter-individual difference, we incorporated an additional analysis in the initial version of the paper that we find compelling in illustrating the relationship between IAF and phase. Specifically, we replicated the binning-based approach, this time binning alpha frequencies not within participants but based on their inter-individual IAF. Specifically, we separated participants into two groups: one with high alpha frequency (average 12.38 Hz) and another with low alpha frequency (average 10.29 Hz). Once again, by avoiding averaging across the entire sample, the differentiation between individuals becomes evident — a difference of 2.09 Hz, which is approximately 42 times greater than 0.05 Hz.

We hypothesized that individuals with higher IAF would sufficiently cover the stimulus, regardless of the specific phase angle at stimulus onset. Faster alpha frequencies naturally provide greater coverage, allowing individuals to adapt even if the stimulus begins in an unfavourable phase. This is also supported by the greater variability characterizing individuals with faster IAF, as empirically demonstrated in the revised manuscript (Fig. 1C). Therefore, if variations in alpha are unable to bring significant differences in the covered phase angle, the role of phase should be less crucial. Conversely, we expected individuals with slower IAF to be more phase-dependent, unable to effectively compensate for starting from an unfavourable phase angle. Our results exactly confirmed this hypothesis: individuals with higher alpha frequencies showed a lesser impact of phase on performance, while those with lower alpha frequencies were significantly influenced by phase. The difference in frequency between these groups (i.e., 2.09 Hz) supports this disparity. For example, individuals with faster alpha cycles (80.77 ms) would cover approximately 74% of an alpha cycle within the 60 ms stimulus window, whereas those with slower alpha cycles (97.18 ms) would cover about 62%. Therefore, even though there is inter-trial and inter-time variability as described in point 1, this has less impact on individuals with a high IAF because, even in the "worst" cases (i.e., 2DS below the mean), a high IAF still allows for sufficient coverage of the stimulus. So, the magnitude of the different phase coverage between the two sub-populations is significant and thus explains the interaction IAF*Phase we robustly demonstrated. In summary, we repeat again that, while the observed differences in IAF may appear small at the grand average level, they are meaningful when considering the broader distribution and variability within individuals. These differences, even if minor in their mean absolute value, can significantly influence behavioral outcomes, supporting the conclusion that alpha frequency influences perceptual sensitivity by modulating the distribution over phases.

3) To provide a broader perspective beyond our own research group, we present 3 figures from highly relevant studies published in top-tier journals (including one in Nature Communications) investigating the relationship between IAF and perceptual outcomes (Nelli et al., Nat. Comm. 2017; Samaha et al., Curr. Biol., 2015; Wutz et al., PNAS, 2018). As it can be clearly seen, the magnitude of differences reported between correct and incorrect responses strictly mirrors what is reported in our findings. This underscores that even subtle differences, which appear subtle only in terms of their mean absolute value for the reasons discussed earlier and which apply equally in these studies, significantly influence perception when averaged across numerous trials and participants.

Nelli et al., Nat Comm, 2017

https://doi.org/10.1038/s41467-017-02176-x

Figure Redacted

Samaha et al., Curr Biol, 2015

https://doi.org/10.1016/j.cub.2015.10.007

Figure Redacted

Wutz et al., PNAS, 2018

https://doi.org/10.1073/pnas.1713318115

Figure Redacted

4) Another very relevant point concerns the statistical interpretation of the findings. A fundamental principle in statistics is that relying on arithmetic or absolute values to assess significance and impact can be misleading. This approach is similar to basing the presence of significance on a figure that represents the average of two measurements. It is important to note that both the analysis highlighted contrasting IAF in correct vs. incorrect trials and the bin-based analysis are statistically compelling, as shown by the Bayes Factor above 50, which strongly supports the hypothesis linking IAF to accuracy.

In conclusion, we hope the reviewer will now convene with us on the statistical strength of our observed effects which indicate their robustness and significance in the relationship between IAF and perceptual sensitivity.

Lines 203-204. There is a statistic report but no indication over which channel/time-point or cluster.

Thank you for the comment. We conducted the highlighted analysis by selecting the IAF from a single electrode for each individual subject. Specifically, we focused on the occipito-parietal electrodes along the midline and right hemisphere (i.e., Oz, POz, O2, PO4, PO8) and, for each subject, we selected from this pool the electrode that exhibited the maximum alpha power during the prestimulus window.

These electrodes were chosen because the stimulus was consistently presented on the left. We have now explicitly reported this in the Methods section to enhance clarity. Additionally, we provided the distribution of selected electrodes as follows: Oz (11%), POz (25%), O2 (11%), PO4 (22%), and PO8 (31%).

Regarding the specific analysis, we acknowledge that we had not previously clarified how we obtained IAF. In the revised manuscript, we have now specified:

> *First, we averaged the IAF across the prestimulus window (-800 to -100 ms) and then applied z-scoring to reduce the impact of outlier trials.*

Given these considerations, the analysis in question was not cluster-based, as we selected a single IAF value per participant (averaged across all time points in the prestimulus window) and one electrode (chosen at the individual subject level).

However, beyond this specific analysis, we also realized that we had not reported the p-values for the cluster-based analyses. In the new version, we have included all p-values derived from the cluster-based analyses throughout the manuscript. Moreover, we described in the Methods section how the p-values were calculated, detailing the statistical approach used for the cluster-based permutation tests:

> *For each cluster, we calculated the p-value by determining the position of the observed cluster within the distribution of clusters generated from the permutation procedure. This approach was consistently applied across all permutation-based analyses.*

Results

Line 337. Integratrion is a typo

We amended the typo.

Lines 358-368. Again, the mechanism proposed is sound and consistent with previous evidences. However, the author should point that within the mechanism they propose these
results should be expected only for stimuli with a specific length.

Thank you for this insightful comment. We have expanded this part of the discussion (see previous comments) in the revised version, incorporating reasoning that links stimulus duration and variability. We believe that the discussion now provides a more comprehensive view of the nature of the effect, offering a clearer framework for how future studies could investigate the modulation of the IAF-phase relationship by short vs. long and dynamic vs. static stimuli.

Figure 3 is missing the letter identifiers (A,B, and C).

Thank you for your comment. We have incorporated the letter identifiers (A, B, and C) in the new version of the figure.

Line 446: "and further ampliated" (while ampliated is indeed a word historically it is rare to find it in regular discourse)

We amended the text using "further extended" instead of ampliated.

Reviewer #3 (Remarks to the Author):

This study presents a novel and interesting view on the mechanisms underlying alpha oscillations' effect on visual perception. They use a visual detection task and investigate how instantaneous and individual alpha frequency and alpha phase impact perceptual performance. The study demonstrates that individual alpha frequency (IAF) significantly influences perceptual accuracy, with faster IAF predicting higher sensitivity and accuracy in sensory acquisition. This relationship is moderated by the phase of alpha oscillations, where individuals with slower IAF exhibit more pronounced phase-related variations in perceptual accuracy, whereas those with faster IAF cover a broader range of phase angles, reducing phase influence on decision accuracy. These findings provide critical insights into previously reported conflicting results regarding alpha phase and perception. The authors used state-of-the-art analysis methods and computational modelling of the behavioral performance. They clearly present the current state of the field and the shortcomings of current theoretical models of alpha oscillations in perception.
I consider that this study is of very high quality and could have an important impact in the field. I have some concerns with the clarity, level of details reported and some of the analytical choices and controls. I list my comments below, roughly in order of importance. I believe that most of my remarks and questions will not change the main conclusions of the paper but hopefully improve its impact and quality.

Thank you for the strong support and overall positive assessment of our work.

1. The DDM analysis is interesting but some aspects should be explained in more details and better motivated:
a. The task's temporal constraints (none in this study) should be indicated before presenting the DDM results as such constraints heavily impacts DDM fitting, e.g. a relatively strict response deadline will impair DDM fit as the RT distribution is truncated. The authors explain how RTs were selected for this analysis in the methods (and their choices seem valid), but providing more information on the task structure (at the beginning of the result section or in the DDM section) would help readers interpret the findings.

We fully agree that the method of reaction time collection is crucial for accurately fitting a DDM model. Imposing a delay between stimulus presentation and participant response could invalidate the model fitting due to truncation of reaction times imposed by the experimental protocol. To avoid this important limitation, we did not impose any delay between stimulus presentation and response. To clarify this and provide a more precise description of the task, we have included a more detailed explanation in the results section:

> *Participants were instructed to indicate via the keyboard the presence (key 'k') or absence (key 'm') of the grey circles inside the checkerboard. No timeout has been set for the response. After collecting the response, the screen appeared black for 1.9–2.4 s in the inter-trial interval.*

b. The authors mention that it takes into account RTs but without elaborating on how RTs are integrated in their proposed mechanistic view of how IAF and phase shape perception. See also my other comment below on the lack of information on RTs in this study.
c. The authors should elaborate on what the DDM analysis adds on top of the SDT analysis as the two indices reported from the DDM (drift and bias) are conceptually very close to the SDT ones (sensitivity and criterion). Or whether the DDM analysis was used as a conceptual replication of the SDT findings.
d. The DDM commonly allows to estimate two other parameters not reported in the current study: bound and non-decision time. These parameters are not reported in the current study. Moreover, in the original paper presenting the DDM model used in the current study, Wiecki et al. (2013) explain that one should to allow all DDM parameters to vary according to conditions (here IAF) and perform model comparisons to determine which parameters should be fixed and which ones should be allowed to vary with conditions. I do not believe that this procedure was followed in the current study.

Thank you for your insightful comments. We agree that further elaboration on these points would strengthen the manuscript. Below, we address each of your concerns in detail.

1) We completely agree that DDM and SDT indices are conceptually related, as both address similar aspects of decision-making. However, there are several key distinctions and advantages to using DDM alongside SDT:

   A. Precision
   DDM provides a more detailed analysis by incorporating both choice and RT data. This allows for a richer understanding of the decision process compared to SDT, which only considers choices. DDM's ability to model how decisions evolve over time offers additional insights into the dynamics of decision-making, especially in a time-sensitive task like ours, where brain temporal dynamics play a crucial role.

   B. Trial-by-Trial Analysis
   The formulation of DDM used (Wieckli et al., 2013) allows for a trial-by-trial analysis of decision processes, which is crucial for our study's aims. This granularity enables us to examine how trial-by-trial fluctuations influence decision-making, something that is not feasible with SDT due to its reliance on aggregated data, which limits its ability to capture these dynamics. Thus, while SDT can indicate general patterns, DDM offers a deeper understanding of the effects at play, demonstrating that the observed outcomes are not merely averages. This approach significantly enhances the reliability of our findings; it mitigates the risk of skewing results due to a small number of trials that might inflate the overall outcomes.

   C. Conceptual Replication
   While there is conceptual overlap between the indices derived from DDM and SDT (e.g., drift rate and sensitivity), using both methods serves as a form of conceptual replication. This approach strengthens our findings by demonstrating that the observed effects are robust across different (and yet related) analytical frameworks.

We have included a more detailed discussion of these points in the method section to clarify how DDM enhances our analysis beyond what is provided by SDT:

DDM offers a more detailed analysis by incorporating both choice and reaction time data, providing richer insights into decision dynamics compared to traditional SDT. Moreover, the DDM formulation we employed allows for trial-by-trial analysis, enabling us to examine the influence of individual trial characteristics on decision-making—something not feasible with SDT's reliance on aggregated data. Additionally, by using both DDM and SDT, we achieve a form of conceptual replication, strengthening our findings across different analytical frameworks.

2) We recognize the importance of following best practices when applying the DDM and agree with the recommendation to allow all parameters to vary according to conditions, as outlined by Wiecki et al. (2013).

Our initial study was guided by a strong top-down hypothesis that specifically focused on the influence of IAF on the drift rate, rather than on other DDM parameters such as boundary separation or non-decision time. We concentrated on drift rate and starting point because these parameters closely align with the concepts of sensitivity and criterion in SDT.

However, in response to your suggestion, we conducted additional analyses using a more comprehensive DDM approach. We fitted a model where drift rate, starting point, boundary separation, and non-decision time were all allowed to vary as a function of IAF. These new analyses revealed that IAF fluctuations specifically impact the drift rate parameter (q-value < 0.001) and boundary separation ($q = 0.03$), while no significant effects were found for the other DDM parameters (all q-values > 0.34). Additionally, the Bayesian Information Criterion (BIC) indicated a better fit for this more complex model compared to the original one. Furthermore, we extended our analysis by investigating whether including all the DDM parameters would improve the fit for the other DDM model discussed in the original manuscript—specifically, the model that examined the roles of IAF, Phase, and their interaction (IAF*Phase). Consistent with our initial findings, the model confirmed that IAF, Phase, and their interaction influenced the drift rate (all q-values < 0.04).

Crucially, the predictors did not have significant effects on the boundary separation and the non-decisional time (all q-values > 0.07), while the Phase impacted the starting point of evidence accumulation: decision bias was more neutral in the positive vs. negative phase (q = 0.02). The goodness-of-fit measures also supported the superiority of these more complex model compared to the previous one. Therefore, in the new version of the manuscript we incorporated the results from these new, more comprehensive models.

3) Regarding the lack of inclusion of RT pointed out in this and subsequent comments, we agree that their inclusion is valuable. For this reason, we replicate the bin-based and the trial-by-trial analysis placing RT, instead of IAF, as dependent variable. The analysis revealed no significant relationship between IAF and reaction time in the task. This finding suggests that, while IAF is predictive of increased response accuracy, it does not generalize to faster response times. We have supplemented our paper with these additional analyses that demonstrate this point, alongside descriptive statistics for RTs, including global RTs, RTs for correct responses, and RTs for incorrect responses. Additionally, we have incorporated a visual representation of the mean accuracy and mean RT in the task in Figure 1.
This result offers a valuable opportunity to further clarify the critical role of the DDM in our analysis, particularly in light of the new formulation you insightfully suggested. Indeed, incorporating all DDM parameters in our analyses—rather than focusing solely on drift rate and starting point—addresses the limitations of more static measures like SDT. While SDT is effective for assessing sensitivity through response accuracy, it does not capture the dynamic aspects of decision-making. Specifically, it is well-known that there is a trade-off between accuracy and reaction time: when accuracy is prioritized, reaction times generally increase, and vice versa.
This trade-off is an essential consideration because a simple increase in sensitivity, as measured by SDT, might not necessarily indicate an improved perceptual performance but could instead reflect a strategic shift toward prioritizing accuracy at the expense of speed. DDM, however, allows us to take this alternative into account by considering parameters like the boundary separation (*a*), which represents the

amount of sensory evidence an individual requires before deciding and non-decisional time (*t*), which represents the time needed for decision-unrelated processes. *a* parameter is crucial because it captures the tendency of a participant to be more cautious (thus more accurate but slower) or more impulsive (faster but potentially less accurate). The fact that IAF significantly predicts the drift rate - even when the boundary separation parameter is accounted for - strongly supports the notion that the relationship between IAF and task performance is not merely a reflection of a shift in the accuracy-RT trade-off. Instead, it indicates an actual improvement in the quality of the perceptual process itself. Higher IAF is not merely linked to a slowdown for the sake of improved accuracy; rather, it is associated with a more efficient accumulation of sensory evidence, leading to better decisions without necessarily requiring more time. The absence of a significant relationship between IAF and RT suggests that IAF is more closely related to perceptual sensitivity rather than the speed of decision-making. In other words, IAF may enhance the system's ability to analyze sensory information, resulting in more accurate responses. However, this improvement in accuracy does not necessarily translate into faster reaction times.

We have integrated these new results and consideration in the main text:

The average RT for the task was 0.88 ± 0.03 s. Correct responses had faster RT (0.85 ± 0.03 seconds) compared to the slower RT for incorrect responses (0.99 ± 0.04 seconds; see Figure 1B).

Moreover, we found no relationship between IAF and RT on a trial-by-trial basis (mean slope = 0.03, SE = 0.10, t_{115} = 0.25, p = 0.80, BF = 0.11), demonstrating that the impact of IAF on accuracy is not due to a speed/accuracy trade-off. If such trade-off were present, higher IAF would be expected to correlate with slower reaction times, as individuals might prioritize accuracy over speed. However, since no such effect on reaction time was observed, it suggests that IAF directly enhances response accuracy without affecting response speed.

And in the supplementary materials:

When examining the relationship between IAF bins and reaction times (RTs), no significant differences were observed between the first and third terciles, whether considered overall ($RT_{\text{first tercile}} = 0.88 \pm 0.03$, $RT_{\text{third tercile}} = 0.88 \pm 0.03$, $t_{115} = -0.19$, $p = 0.85$) or when weighted by response accuracy ($RT_{\text{first tercile correct}} = 0.85 \pm 0.03$, $RT_{\text{first tercile incorrect}} = 1.00 \pm 0.04$; $RT_{\text{third tercile correct}} = 0.86 \pm 0.03$, $RT_{\text{third tercile incorrect}} = 0.97 \pm 0.04$, all $t_{115} < 1.42$, all $p > 0.16$). We also tested whether IAF fluctuations were associated with RT in the task, demonstrating that variations in IAF did not significantly predict RT on a trial-by-trial basis (mean slope = 0.05, SE = 0.10, $t_{115} = 0.52$, $p = 0.60$, BF = 0.12). These results collectively support the notion that IAF has a significant impact on the probability of accurate perceptual responses while leaving reaction time unaffected. This finding is important as it indicates that the impact of IAF on accuracy is not due to a speed/accuracy trade-off. If such a trade-off were present, higher IAF would be expected to correlate with slower reaction times, as individuals might prioritize accuracy over speed. However, since no such effect on reaction time was observed, it suggests that IAF directly enhances response accuracy without affecting response speed.

e. The authors do not describe how the diagram representing the DDM model (Fig. 2C) was obtained. Is it a generic graphical representation of the DDM or were the distributions and drift arrows simulated based on the model fit to the data? Also, the diagram should be described in more details in the caption.

Thank you for pointing this out. The diagram representing the DDM model in the figure shows the traces of the actual model fitted to the data, rather than being a generic graphical representation or simulated output. This detail was described in the Methods section where it is stated that the traces were plotted based on the fitted model.

To be crystal-clear on this point, we updated the figure caption of the figure in the new version of the manuscript to include this information, ensuring that it clearly conveys that the diagram reflects the output of the fitted model.

2. The behavioral task is not described in sufficient details:
a. The stimulus composition is unclear. If I understand correctly, every stimulus was a checkerboard composed of black and white cells in catch trials, and with superimposed grey circles in target trials. Was a target composed of the checkerboard with a circle in every cell or only a subset of cells? It seems like it was in every cell but this is not explicitly mentioned in the methods.
b. What was the size of the entire stimulus (in degrees of visual angle)? What was the size of each square/what was the spatial frequency?
c. In the only depiction of the stimuli in Fig. 5, it seems like circles placed on white versus black cells had different shades of grey, what was the difference between these circles' colors? What specific feature of the circles was titrated in the staircase?
d. The staircase procedure is not explained, please provide more information on the type of procedure used (1-up-2-down? Quest method?), how many trials were included in this staircase, etc.

Thank you for your feedback. We appreciate the opportunity to clarify the methods and stimulus details used in our study.

a. Stimulus Composition:

> The stimulus composition was indeed described in the Methods section, but to ensure clarity: in target trials, each checkerboard stimulus featured grey circles in every cell. In contrast, catch trials consisted of checkerboards with no grey circles. This distinction was explicitly stated, but we have now added further details to ensure this is clear (see below).

b. Stimulus Size and Spatial Frequency:

We have revised the Methods section to include the following specifics about the stimulus:

Stimuli had a spatial frequency of 5.16 cycles/degree, and they were presented only in the lower part of the screen at 4.1°/3.7° eccentricity (horizontal/vertical).

c. Circle Colors:

We acknowledge that the information concerning the grey stimuli was not sufficiently detailed. Regarding the color difference observed in Figure 5 (now Figure 7), although the circles on white and black cells have different shades of grey, the images themselves were designed to have identical average luminance across conditions. Specifically, the contrast level of the grey circles was systematically varied in relation to the background to maintain this luminance consistency. For clarity, let's consider some practical examples: the checkerboard cells background could either be black (RGB value 0, 0, 0) or white (RGB value 255, 255, 255). The contrast of the grey circles embedded in these cells was adjusted based on the background color. For instance, in the case of the image with RGB values of 55/200, the grey circle superimposed on the black cells had an RGB value of 55, 55, 55. In contrast, the grey circle superimposed on the white cells had an RGB value of 200, 200, 200.
In other images, the contrast was further manipulated, bringing the grey values closer or farther from the background RGB values, thus altering the difficulty of detection. For instance, the image with RGB values of 25/230 featured grey values that were significantly closer to the background value, thereby making detection more challenging. In contrast, the image with RGB values of 75/180 presented grey values that were much farther from the background value, facilitating easier detection. This contrast was systematically varied to determine the perceptual threshold at the individual level. Moreover, by definition, the catch trials consisted solely of the background, meaning the RGB values could either be 0, 0, 0 (for black cells) or 255, 255, 255 (for white cells). Importantly, this strategy ensured that the

average RGB level of the targets and catch stimulus—and therefore their overall luminance—remained constant across all selected image levels.

d. Staircase Procedure Details:

We acknowledge that the staircase procedure was not sufficiently detailed. To address this, we have provided a more comprehensive explanation in the Methods section. Furthermore, while we have published previous works where this approach was employed (Tarasi and Romei, 2024? JOCN) we would like to highlight that a systematic validation of this threshold estimation strategy is currently under review in a separate article, which offers a more extensive validation compared to classical methods.

Following a training phase, each participant underwent a staircase procedure to determine the contrast level of the gray circles needed to achieve 70% detection accuracy. This procedure was conducted with an equal number of target-present and target-absent trials. Specifically, participants completed blocks consisting of five trials with targets present and five trials with targets absent. After each block, the participant's accuracy was assessed. If the accuracy was 50% or lower, the gray contrast was increased by 4 points. If the accuracy was below 70%, the contrast was increased by 1 point. Conversely, if the accuracy was 90% or higher, the contrast was decreased by 4 points, and if it was above 70%, it was decreased by 1 point. This procedure was repeated 25 times. The stimuli used included checkerboards with gray circles embedded, with the lowest contrast being red, green, and blue (RGB) values of 15/240 (where the first value refers to the contrast of the circles against the white squares and the second value against the black squares) and the highest contrast being 75/180. The initial contrast for threshold estimation was set between these two extremes (RGB value: 45/210). The detection threshold was estimated based on the final three blocks (Figure 1C). Including target-absent trials was crucial for obtaining a bias-free measure, avoiding confounding effects related to participants' decision criteria. Without target-absent trials, it would be difficult to determine if variations in threshold values were due to *actual perceptual ability or simply the number of hit rates, despite equal sensitivity when controlling for false alarms (Green & Swets, 1966). This approach aligns with the Signal Detection Theory (SDT) principle, where d' is calculated by considering both hit rates and false alarms to derive a bias-free measure of performance. The bias-free approach means that, assuming equal sensitivity, the choice of a liberal or conservative decision criterion does not affect the estimation of the contrast threshold, unlike classical staircase methods (e.g., two-down, one-up methods; Levitt, 1971). Notably, this threshold estimation method is novel and was first implemented in Tarasi et al. (2022) and in Tarasi and Romei (2023).*

e. No information is provided on reaction times in this task (neither descriptive statistics nor any statistical test) although this performance index is used in the DDM analysis. This information is essential to evaluate the paradigm in general and some of the findings, and should thus be presented at the beginning of the results section (at least descriptive stats and/or graphical representation).

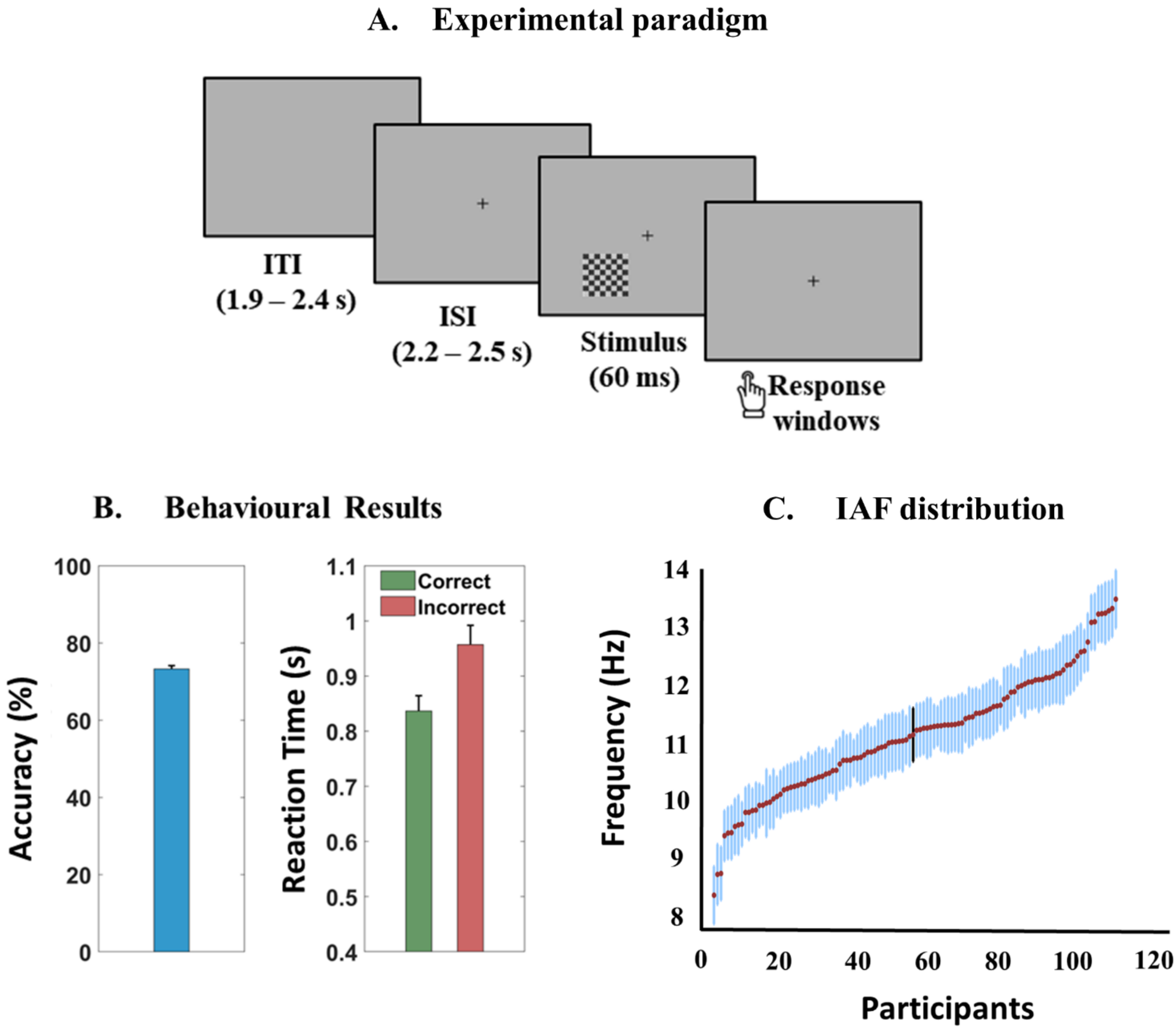
As we stated in a previous comment, we have now supplemented our paper with additional analyses that demonstrate the relationship between IAF and RT, alongside comprehensive descriptive statistics for RTs, including global RTs, RTs for correct responses, and RTs for incorrect responses. Additionally, we have incorporated the main accuracy metrics as well as the mean reaction times for both correct and incorrect decisions in a new figure included in the new version of the manuscript (Figure 1B).

3. In general, I think the task should be explained in more details at the beginning of the results section and accompanied with graphical depiction of the trials time course, and together with descriptive statistics (and/or plots) of accuracy and RTs.

We have taken your suggestions into account and revised the text to enhance the clarity and comprehensiveness of the experimental task. We have now provided a more explicit and detailed description of the experimental protocol at the beginning of the Results section. This revision ensures that readers have a clear understanding of the task before reading the actual results:

> *Each checkerboard was presented for 60 ms and the contrast of the grey circles was titrated for each individual to their 70% perceptual accuracy. In half of the trials, checkerboards did not contain any gray circle embedded in them (catch trials).* ***Participants were instructed to indicate via the keyboard the presence (key 'k') or absence (key 'm') of the grey circles inside the checkerboard. No timeout has been set for the response. After collecting the response, the screen appeared black for 1.9–2.4 s in the inter-trial interval.***

Moreover, we have included a new figure (Figure 1A, see below) that graphically depicts the time course of the trials, offering a visual representation of the task structure together with descriptive plots of mean accuracy and RTs (Figure 1B).



4. Because (one of) the main axe(s) of the study is individual differences in IAF, the authors should document this further.

a. It would be informative to know how much IAF varies within an individual and whether this variability is correlated with the individual's average IAF (e.g. is prestimulus IAF more/less variable in individuals with higher average IAF?). One inspiration for representing this variability could be Fig. 2 of Grabot & van Wassenhove (2017) (https://journals.sagepub.com/doi/full/10.1177/0956797616689369?journalCode=pssa) investigating interindividual differences in temporal-order-judgments. Of course, this is just a suggestion for graphically representing this variability.

Thank you for this valuable feedback. We did appreciate your suggestion regarding the visualization of individual differences in IAF variability. In response, we have added a new figure (see the previously attached Figure 1C) to the manuscript. This figure illustrates the variability in alpha oscillations on a subject-by-subject basis across trials. Moreover, we conducted a correlation analysis to assess whether differences in IAF were reflected in different variability of alpha. The results revealed an interesting relationship: individuals with higher pre-stimulus IAF also exhibited greater alpha variability (Pearson = 0.37, $p < 0.01$). Further studies could investigate how variability and IAF relate to each other and their combined impact on behavior, potentially examining whether individuals with higher IAF are better able to modulate their IAF by slowing down or increasing it as required by the context (Wutz et al., 2018). Additionally, the newly included phase analysis (see comment below) clearly demonstrates how this variability in turn influences the relationship between IAF and phase. Specifically, individuals with a lower IAF appear more rigidly tied to the stochastically present phase, leading to a stronger clustering of phase angles associated with correct vs. incorrect responses. This effect arises because being positioned at disadvantageous angles in the prestimulus period increases the likelihood of errors due to reduced phase coverage. Conversely, a faster and more variable IAF, as shown in this analysis, allows for sampling across a wider range of phase angles within the same timeframe. This expanded sampling increases the chances that a disadvantageous phase position at time x can be compensated at time x + t, thereby reducing the likelihood of a distinct clustering of phase angles between correct and incorrect responses. We chose not to elaborate further on this point, as it goes somewhat beyond the scope of our work and could add unnecessary complexity that might detract from the main focus. However, given its intriguing nature, we have included this observation in the discussion:

> *But how can IAF and alpha phase interact? The core reasoning relies on the assumption that an individual with a higher alpha frequency can traverse a broader range of phase angles within the same temporal interval compared to an individual with a lower alpha frequency (Figure 5). To illustrate, envision two individuals, one showing a relatively higher IAF and the other showing lower IAF, processing stimuli within the same timeframe. The one with higher alpha covers more "ground",*

encountering and embracing a multitude of phase angles. In contrast, the lower alpha counterpart covers a narrower spectrum of phase angles during the identical processing window. This distinction becomes pivotal in understanding decision outcomes, especially in individuals with lower IAF. For them, the positioning of the alpha phase at specific time points becomes a critical determinant: for individuals with lower alpha, the transition from the unfavourable phase bin to the favourable one precisely at those crucial points is less likely. In essence, the restricted coverage of the lower alpha diminishes the probability of aligning with optimal phases during stimulus presentation, intensifying the impact of phase positioning on decision outcomes. ***Moreover, this effect would be amplified by the greater IAF variability observed in individuals with higher IAF during the prestimulus period. This increased variability enhances their advantage, as it allows for more dynamic phase coverage, maximizing the likelihood of aligning with favorable phases during task performance and thereby optimizing decision accuracy.***

b. Moreover, since objective target discriminability was titrated for each participant, and according to the authors' claims, there should be a correlation between participants' average IAF and their titrated contrast reached during the staircase. Showing this correlation would strengthen the authors claim on the impact of IAF on perceptual performance.

Thank you for your comment. We appreciate the suggestion to explore the correlation between participants' average IAF and the titrated contrast levels reached during the staircase procedure.

We would like to emphasize that a relationship between IAF and contrast levels has been demonstrated in a previous and recent study of ours, published in the *Journal of Cognitive Neuroscience* (https://doi.org/10.1162/jocn_a_02026). In that study, we observed a significant correlation indicating that individuals with a higher IAF required lower contrast levels to achieve accurate performance. This finding aligns with our current hypothesis and

provides robust evidence for the impact of IAF on perceptual performance. We have made sure to clearly incorporate this finding in the revised manuscript to further support our claim about the impact of IAF on perceptual performance. Furthermore, prompted by your comment, we conducted a similar analysis correlating mean pre-stimulus IAF with individual contrast thresholds. The results closely align with those from our analysis of resting-state IAF: individuals with higher IAF exhibited lower contrast thresholds (Pearson $r = -0.21$, $p = 0.03$; Pearson skipped $r = -0.21$, CI = [-0.37, -0.03]). We have incorporated this additional evidence linking IAF to perceptual sensitivity in the revised version of the paper.
Specifically, in the introduction we acknowledged the previous finding:

> *Additionally, we have recently demonstrated that individuals with a higher resting IAF require less contrast for accurate detection compared to low IAF participants (Tarasi and Romei, 2023).*

While in the results we integrated the new finding:

> *Building on our recent finding that individuals with higher resting IAF require less contrast for accurate detection compared to those with lower IAF*(Tarasi and Romei, 2023)*, we further demonstrate that this relationship extends to average pre-stimulus alpha levels. Specifically, individuals with higher pre-stimulus IAF consistently required lower threshold contrast to achieve target performance levels (Pearson = -0.21, p = 0.03; Pearson skipped = -0.21, CI = [-0.37, -0.03]).*

5. Some details are missing in the EEG methods:

a. I guess the authors only inspected occipito-parietal electrodes on the midline and right side because the stimuli were always displayed on the left visual hemifield. This should be explicitly stated in the methods section.

Thank you for pointing out this important detail we missed to explicitly mention in the result included in the first version of the manuscript. In the current version we stated that:

> *We focused our analysis on a pool of occipito-parietal electrodes located on the midline and right side (i.e., Oz, POz, O2, PO4, PO8), selecting, for each subject, the electrode that exhibited the maximum alpha power in the prestimulus window. These electrodes were chosen because the stimulus was consistently presented on the left.*

b. The authors omit important details on the EEG recording, namely: how many electrodes were used and what system/brand/device was used?
c. In the ICA preprocessing step: how many components were the data decomposed in? How many components were discarded per participant on average? The author mention that "components containing artifacts that could be clearly distinguished from brain-driven EEG signals", does that mean that all "non-brain-driven" component were discarded? In a typical 64 electrodes EEG dataset ICA decomposition, if the ICA decomposes the signal into 64 components, most components are not brain-driven, therefore it would mean that in this study a large number of components were discarded. Or did the author discard a specific and restricted type of components (e.g. blinks, lateral eye-movements, as is commonly done)?

Thank you for pointing this out; we acknowledge that our initial description was somewhat general. We have used a Brain Products actiCHamp EEG amplifier configured with 64 EEG channels, mounted according to the international 10–10 system. The number of EEG components extracted was equal to the number of electrodes. On average, 6.4 ± 0.38 components were removed. For the removal, we primarily focused on components associated with eye movements, blinks, muscle artifacts, and any noisy electrodes.

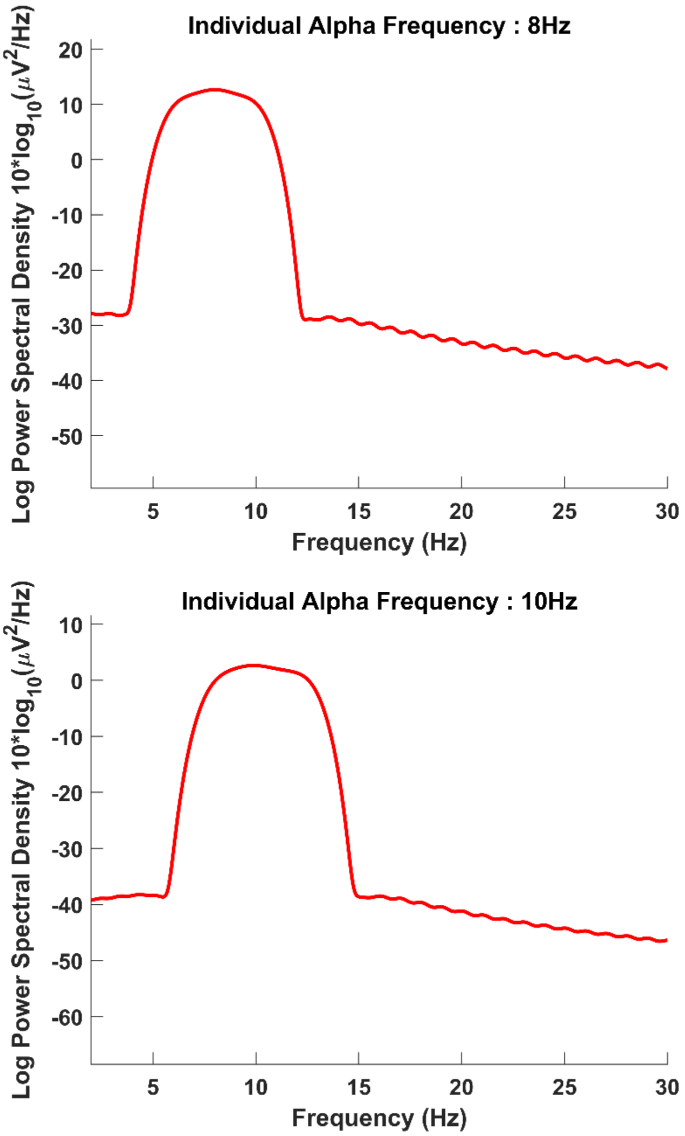
We have included this information in the revised version of the manuscript:

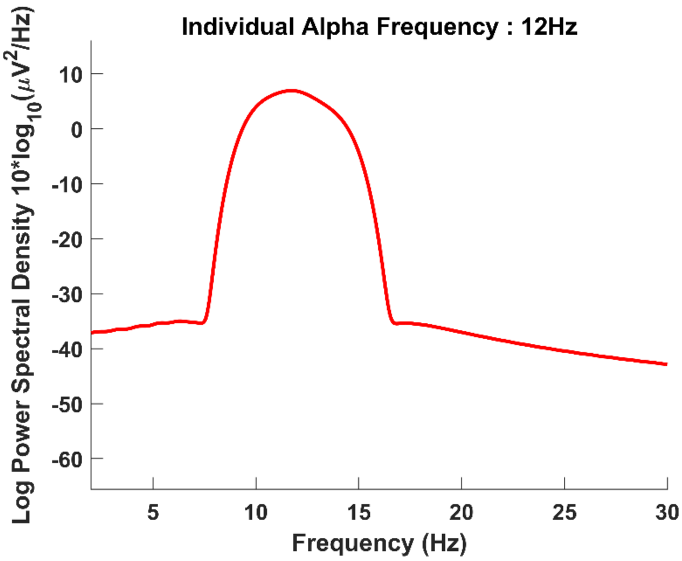
> ***We used a Brain Products actiCHamp EEG amplifier with 64 channels, configured according to the international 10–10 system.*** *EEG was measured*

with respect to FCz electrodes, and all impedances were kept below 10 kΩ. EEG signals were acquired at a rate of 1000 Hz. EEG was processed offline with custom MATLAB scripts (version R2021a) and with the EEGLAB toolbox(Delorme and Makeig, 2004)*. The EEG recording was filtered in the 0.5–100 Hz band and a notch-filter at 50 Hz was applied. The signals were visually inspected, and noisy channels were spherically interpolated. Then, trials containing excessive noise, muscle or ocular artefacts was discarded. An average of 149 trials per subject passed this stage. Next the recording was re-referenced to the average of all electrodes, and the Independent Component Analysis (ICA) was applied.* ***On average, 6.4 ± 0.38 components were removed, primarily those associated with eye movements, blinks, muscle artifacts, and noisy electrodes.*** *The signal was downsampled to 256 Hz*.

d. The filtering should be described in more details (type of filter, cut-off frequencies, etc.) as this could potentially have an impact on oscillatory measures and would be necessary to replicate this study.

We agree with the reviewer about the importance of designing an appropriate filter for extracting the instantaneous alpha frequency. As included in a similar comment raised by reviewer 1, upon review, we found that the initial filters employed slightly different parameters compared to those commonly used in the literature (London et al., 2022), so we adjusted them accordingly and reported the details about it. Specifically, we filtered the signal using a zero-phase, plateau-shaped, band-pass filter with 15% transition zones (Matlab function *filtfilt.m*). Moreover, we calibrated the cut-off frequencies at the individual subject level, setting them from IAF - 2 to IAF + 2. This choice is based on recent methodological work by Samaha and Cohen (2022), which demonstrated that centering the filter around the IAF is crucial for obtaining reliable estimates of instantaneous frequency. It is important to highlight that the overall pattern of results remains fully consistent with the initial version. If anything, the modulation has further strengthened the previous findings. Attached are some figures showing the power spectrum obtained using the *spectopo* function after the filtering process for selected subjects. As observed, the filter effectively captures the alpha band and crucially adapts to the precise range based on each subject's individual alpha frequency, while frequencies outside this range are attenuated.





We have included in the new version of the manuscript more detail about filter characteristics:

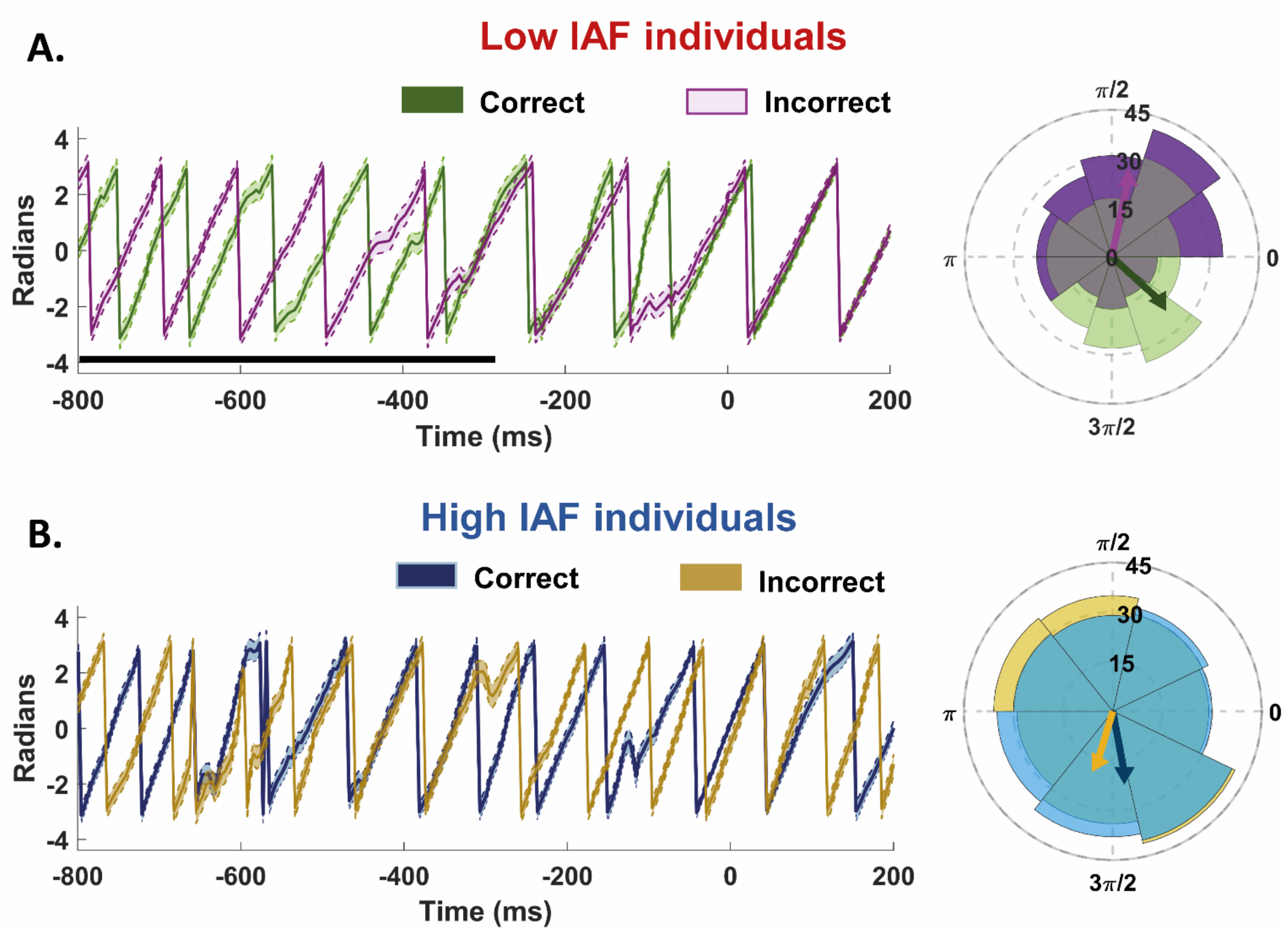
> *we filtered the signal with a zero-phase, plateau-shaped, band-pass filter with 15% transition zones (Matlab function filtfilt.m) centered on the individual oscillation peak extracted in the previous stage (filter range from IAF-2 Hz to IAF+2 Hz.*

6. For the analyses described in the "Instantaneous alpha phase shapes perceptual accuracy" section, the authors use 2 phase bins and argue that this number of bins was chosen to have sufficient number of trials per bin. I understand aiming to retain enough trials per bin but I don't understand how the two bins were chosen. I am not an expert on phase analyses but since absolute phase of neural oscillations measured with EEG is meaningless (peaks and throughs depend on cortical folding and local or distant orientation of the dipoles contributing to the signal), wouldn't it be more sensitive/statistically powerful to choose phase bins individually using a data driven method? For instance, computing a phase opposition index between correct and incorrect trials? I might be missing a critical aspect of the analysis that would not be possible with another method, but I would like the author to better motivate this analytical choice.

Thank you for your insightful feedback. To provide a more comprehensive overview, we included new analyses examining the phase-accuracy relationship, and we also extended the phase analysis window to cover the time period during which we examined the IAF-Phase relationship. This decision was motivated both by the type of analysis introduced—based on angular difference analysis between two response categories (in our case, correct vs. incorrect), as in the influential paper by Samaha et al., PNAS 2015, which used wider windows—and by the need to highlight phase-related phenomena that unfold over different temporal periods. Expanding the analysis window allowed us to connect these phenomena to the relationship between IAF and variability, which, following your suggestion, we empirically confirmed.

First, we describe the results of the new analysis and how it addresses your request for a more data-driven approach. We then present the findings from the reanalysis of those included in the first version, in which we broadened the windows of interest to ensure a consistent approach. Finally, we discuss the overall results to provide a unified interpretation of the IAF-phase relationship.

In the new included analysis, we compared the mean phase angles for accurate and inaccurate responses using the Watson–Williams test, a circular analog of the t-test that tests the hypothesis that two angle samples have distinct phase distributions (Berens, 2009). This test accounts for both the mean phase angle and the circular variance of each sample and is widely used to assess phase differences electrophysiological data(Arnal and Giraud, 2012; Samaha et al., 2015; VanRullen, 2016; Rideaux et al., 2023). We conducted this analysis separately for each time point within the -800 ms to 200 ms window. Finally, we performed a cluster-based permutation test, evaluating clusters based on the number of consecutive time points in which the distribution of phase angles was different between accurate and inaccurate responses to determine whether they exceeded the threshold derived from the permutation test. We present here the graphical representation of the finding derived from the analysis:



As can be observed, there is an extended temporal period where phase angles differ between correct and incorrect responses. Crucially, this effect is present in the group of individuals with low IAF, while no significant cluster of angle differences is observable in the high IAF group. We would like to emphasize that this analysis does not involve any a priori clustering of phase angles. By considering the mean phase angles in correct vs. incorrect conditions, it is possible to assess in a data-driven manner whether the angles differ significantly within the sample without imposing any constraints, as the continuous angle values from -π to +π are taken into account.

We reported here the description of the analysis included in the methods:

Watson-Williams test for phase difference between correct and incorrect choices

> *To investigate whether accurate and inaccurate perceptual decisions are associated with different phase angles, we analyzed phase data across timepoints ranging from -800 to 200 ms. Specifically, we aimed to determine if the phase angles differed as a function of decision accuracy using the Watson-Williams test. This test, the circular analog of the t-test, evaluates whether two samples of phase angles (corresponding to correct vs. incorrect decisions) are drawn from different distributions (Berens, CircStat: A MATLAB toolbox for circular statistics, 2009). The Watson-Williams test accounts for both the mean phase angle and the circular variance within each sample, making it well-suited for analyzing phase differences in electrophysiological data (Arnal et al., 2014; Busch et al., 2009; Samaha et al., 2015; David et al., 2023; Rideaux et al., 2023). To assess the significance of phase differences, we employed a permutation-based analysis. For each timepoint where the Watson-Williams test indicated a significant difference in phase angles between correct and incorrect decisions, we calculated the corresponding F-value and summed these values across contiguous significant timepoints. We then compared the observed clusters of F-values with a null distribution of clusters generated by permuting the association between phase angles and decision outcomes 1000 times (i.e., shuffling the phase-accuracy relationship). This approach allowed us to determine whether the observed clusters were significantly larger than those expected under the null hypothesis. The analysis was performed on the entire sample, as well as separately for the low and high IAF groups, to explore potential differences between these groups.*

And what we added in the results section:

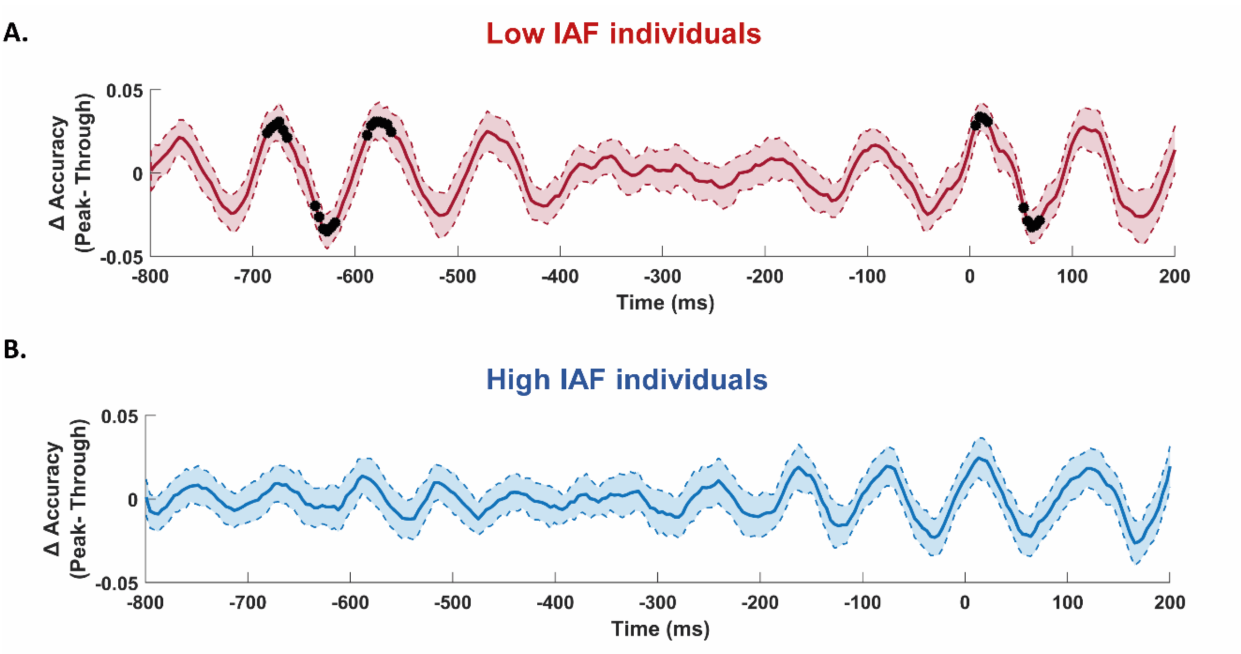
> *To further validate our findings beyond the bin-based approach on phase angles, we employed a reverse data-driven method to examine whether correct versus incorrect responses were associated with distinct phase angles. Watson-Williams test revealed that, while no significant clusters emerged in either the whole sample or the high IAF group (Fig. 5B, the low IAF group showed a notable cluster spanning from approximately -800 ms to -300 ms (Fig. 5A). Within this interval, the phase*

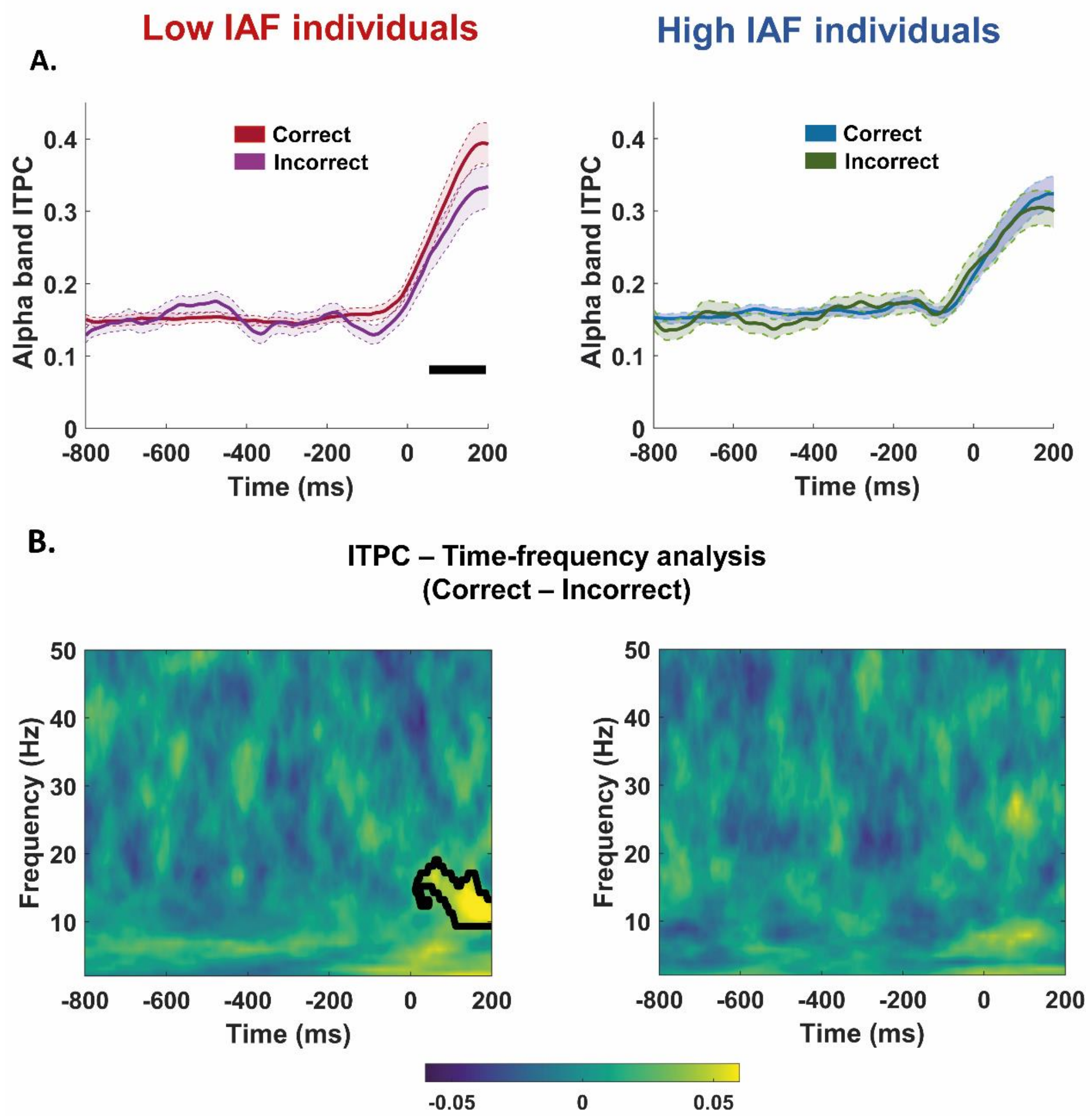
> *angles associated with correct and incorrect responses significantly differed (p = 0.01).*

And what we added in the discussion:

> *Supporting this finding, a complementary analysis—investigating whether the mean phase angle differed between correct and incorrect responses—uncovered a significant, group-specific effect. Specifically, only in the low IAF group did the phase angle differ significantly between correct and incorrect decisions.*

Furthermore, as previously mentioned, we have also re-ran all phase-related time-resolved analyses included in the first version, now considering the extended interval from -800 to 200 ms. As shown in the graph below, the results remained highly consistent:





The phase analysis assessing accuracy based on binning the phase angle into two extremes showed the same pattern of results as in the previous version. Additionally, in the prestimulus period around -600 ms to -500 ms, temporal periods emerged where the different phase bins could predict response accuracy in the low IAF group. Notably, the prestimulus time period in which the phase bin could predict accuracy aligns with the one identified by the phase comparison analysis using the Watson–Williams test. Therefore, both the top-down binning approach and the data-driven analysis—where we assessed whether phases differed based on correct vs. incorrect responses—show a consistent pattern of results in the prestimulus period. Again, no effect at any time period was observed in the high-IAF participant group. Similarly, ITPC analysis showed the same pattern when considering the new time period. Only participants with low IAF were characterized by stronger phase coherence in correct vs. incorrect responses, whereas the high IAF group showed no difference in ITPC.

Overall, the analyses consistently indicate that phase impacts cognitive performance, with this effect modulated by individual differences in alpha speed. The new analysis, examining point-by-point phase differences between correct and incorrect responses, along with the reprocessed previous analyses, reveals a phase-related effect in individuals with low IAF. By expanding the time window, we gain further insight into this phenomenon: specifically, the Watson-Williams analysis unveiled that, in an extended period from -800 ms to -300 ms, phase angles differentiate between correct and incorrect responses only in the low-IAF group. Within this interval, the phase pattern in the low-IAF group is quite rigid, displaying a clear separation of phase angles between correct and incorrect responses. This rigidity suggests that a slower system, which cannot cover multiple phase angles, lacks the flexibility to adapt to potentially unfavourable phases and instead remains consistently locked to the stochastically determined phase.

In contrast, the high-IAF group displays a more variable phase pattern, with phase angles that lack a robust and significant distinction between correct and incorrect responses. This variability likely reflects the system's greater capacity to cover multiple phase angles, reducing the impact of any single phase angle on performance. With more opportunities to engage favourable phases across the analyzed period, the high-IAF system shows greater robustness to unfavourable phase. This result implies a fundamental difference in system variability between the two groups. Specifically, the data indicate that while the low-IAF group adheres to a more rigid pattern—where phase angles for correct and incorrect responses remain largely separated—the high-IAF group demonstrates significantly greater variability, with correct and incorrect phase angles intermixing at multiple time points.

These findings align with previous observations of a relationship between IAF and variability, providing concrete evidence that a faster oscillatory speed enables the system to engage a broader range of phase angles. This flexibility enhances the system's ability to counterbalance potentially unfavourable phases due to the increased variability inherent in higher IAF.

7. The authors used 3 phase bins in the IAF analysis. They mention using the same procedure as another study but they should provide some evidence that the results do not depend on this specific setting. Do the result hold (at least qualitatively) when using 2 or 4 frequency bins? Moreover, the authors only compare the 1st and 3rd bin, I think this is a valid analytical choice, but the effect in the 2nd bin should be presented at least qualitatively or graphically. It would be valuable to know whether the effects observed in the 2nd bin are in-between the 1st and 3rd, or whether there is a non-linear relationship between IAF and perceptual performance.

Thank you for the comment, we appreciate the opportunity to clarify and expand on the analysis concerning the use of phase bins in our study.
First, we have conducted an ANOVA on accuracy and d', and criterion, which includes the second bin as per your suggestion. The ANOVA results for accuracy and d' were significant, while the criterion was not. A one-tailed post-hoc analysis (selected due to the corroborative nature of this analysis, where we have a clear expectation of the direction of the results) indicated significantly greater accuracy and d' in the third bin compared to the first and second bins, while there was no significant difference between the first and second bins. Moreover, the graph included in the supplementary materials does demonstrate a monotonic increase in accuracy from the first to the third bin. To further substantiate this observation, we fitted a regression line to the accuracy and d' data across the three bins for each participant and calculated the slope. We found a significant and positive slope for both measures, reinforcing the notion of a gradual increase in performance across the bins.
Lastly, we replicated the bin-based analysis with different divisions to ensure that our results do not depend on the specific three-bin division. Specifically, we employed a median-split method, which divides the trials into two bins. This approach revealed that both accuracy and d' were higher in the above-median IAF bin compared to the below-median IAF bin. This finding suggests that the observed effects are robust and do not

depend on the specific number of bins used in the analysis. Results do not change when using a division in quartiles, as specified in the revised manuscript. We have integrated this evidence in the new version of the manuscript.

Specifically, in the method section we wrote:

> *To ensure the robustness of our findings, we conducted additional analyses to verify the consistency of our results across different bin configurations. Initially, we expanded our primary analysis to include the second IAF bin and performed ANOVA on accuracy, d', and criterion measures across the three IAF bins. This approach allowed us to assess whether the differences observed remained significant when including the second bin, rather than just focusing on comparisons between the first and third bins. In addition, we conducted regression analyses to assess whether there was a monotonic trend in accuracy and d' across the three bins, which would suggest a gradual increase in performance with higher IAF (see Supplementary Materials for details). To further validate our findings, we also explored the impact of different bin configurations. Specifically, we divided the trials using two and four bins. For the two-bin configuration, trials were divided into above-median and below-median IAF bins. A paired t-test was used to compare accuracy and d' between these two trials type. For the four-bin configuration, we divided trials into quartiles and assessed performance measures across these quartiles.*

And in the supplementary materials we included these analyses:

> *We expanded our primary analysis by using a median-split analysis to divide the trials in two-bins. The pattern of results did not change with this new bin configuration as indicated by the significant differences in accuracy (accuracy $_{above\text{-}median\ bin}$ = 0.75 ± 0.01, accuracy $_{below\text{-}median\ bin}$ = 0.72 ± 0.01, $t_{115} = -3.71$, $p < 0.01$) and d' (d' $_{above\text{-}median\ bin}$ = 1.65 ± 0.07, d' $_{below\text{-}median\ bin}$ = 1.46 ± 0.07, $t_{115} = -3.10$, $p < 0.01$) between above-median and below-median IAF bins.*
> *The results were further corroborated when we used a quartile approach to bin the data. One way ANOVA showed a significant effect of the quartile on accuracy (F*

$_{3,354}$ = 8.84, $p < 0.001$) and on d' ($F_{3,354}$ = 6.72, $p < 0.01$). Post hoc analysis using one-tail paired ttest showed that accuracy in the first quartile (accuracy $_{first\ quartile}$ = 0.72 ± 0.01) and in the second quartile (accuracy $_{second\ quartile}$ = 0.73 ± 0.01) were statistically different from the accuracy computed when considering the third (accuracy $_{third\ quartile}$ = 0.75 ± 0.01) and the fourth quartile (accuracy fourth $_{quartile}$ = 0.75 ± 0.01, all $t_{115} > 3.14$, all $p < 0.01$). Other comparisons did not reach the significance level (all $t_{115} < 1.2$, all $p > 0.23$). Similarly, d' in the first quartile (d $_{first\ quartile}$ = 1.39 ± 0.07) and in the second quartile (d $_{second\ quartile}$ = 1.37 ± 0.07) were statistically different from the d' computed when considering the third (d third $_{quartile}$ = 1.61 ± 0.07) and the fourth quartile (d fourth $_{quartile}$ = 1.51 ± 0.07, all $t_{115} > 2.39$, all $p < 0.019$). Other comparisons did not reach significance (all $t_{115} < 1.45$, all $p > 0.15$).

Moreover, in the supplementary materials, we included an ANOVA analysis that considers all three bins (i.e., including the second bin). We also provided an analysis demonstrating that the slope of the regression line for accuracy and d' across the three bins is positive and significantly different from zero. This confirms that the relationship is monotonic, with performance increasing as a function of IAF frequency:

We conducted an ANOVA comparing accuracy, d' and criterion across the second phase bin along with the first and third bins. Significant values were obtained when considering accuracy ($F_{2,230}$ = 7.22, $p < 0.001$) and d' ($F_{2,230}$ = 5.17, $p < 0.01$), but not criterion ($F_{2,230}$ = 2.13, p = 0.13). These results testify that a significant difference exists between the three terciles. Post hoc analyses using one-tail paired ttest revealed significant differences in accuracy and d' between the first and third bins (as reported in the main text). Regarding accuracy, there was also a difference between the first and the second bin (t_{115} = -1.87, p = 0.03) and between the second and the third bin (t_{115} = -1.92, p = 0.03) Moreover, a significant difference emerged between the second and the third bins when considering d' (t_{115} = -1.92, p = 0.02) but no differences when comparing the first and the second bin (t_{115} = -1.32, p = 0.09).

Additionally, we provide regression analyses that demonstrate a significant positive slope for accuracy (mean slope = 0.016 ± 0.001, t_{115} = 3.71, $p < 0.01$) and d' (mean *slope = 0.09 ± 0.03, t_{115} = 3.09, p < 0.01) but not for criterion (mean slope = -0.02 ± 0.001, t_{115} = 1.67, p = 0.10) across the three bins, reinforcing the notion of a gradual and monotonic increase in performance with higher IAF.*

8. The authors do not seem to control for alpha power in any of the analyses. Although the effect of alpha power on perceptual processes is still debated in the literature, this could be an important confound in the reported analyses, e.g. alpha power in trials grouped in the 1st and 3rd bins according to IAF values could be different and thus put the authors' claims into question. This should also be investigated at the interindividual level: is alpha power different in low versus high IAF participants?

We fully agree with the reviewer on the importance of clarifying the specificity of the effects linking IAF to perceptual performance, ruling out the possibility that the same relationship could be driven by alpha power. To address this concern, we conducted a series of analyses that consistently demonstrated that alpha power does not significantly influence perceptual performance. Below is a summary of the analyses, which are detailed in the supplementary materials.

First, we extracted time-frequency amplitude maps and conducted a cluster-based permutation test to assess whether oscillatory amplitude differed based on response accuracy (correct vs. incorrect trials). The statistical analysis revealed no significant clusters distinguishing correct from incorrect responses (p = 0.43).

Next, we performed regression analyses using a least-squares approach, adding alpha power as an additional predictor alongside IAF. The results indicated that IAF remained a strong predictor of performance (slope = 0.37, SEM = 0.09; t_{115} = 4.04, p < 0.001, BF = 179.63), while alpha power (mean slope = -0.14, SEM = 0.09; t_{115} = -1.57, p = 0.12, BF = 0.34) did not demonstrate any predictive value. A similar pattern emerged when using a GLM model (mean slope for alpha power = -0.03, p = 0.13; mean slope for IAF = 0.08, p < 0.001).

We also conducted bin analyses by dividing trials into terciles based on alpha amplitude. The results showed no significant differences in sensitivity and criterion between trials in the first and third terciles (all p > 0.09, all BF < 0.41).

Finally, we conducted another time-frequency analysis comparing oscillatory amplitude between participants in the slow and fast IAF groups. No significant clusters were found between the two groups (p = 0.14).
In conclusion, these analyses consistently demonstrate that alpha power does not contribute to explaining perceptual performance, in contrast to the robust influence of IAF.

Here we report the new section added in the methods section:

Assessing the specificity of the relationship between IAF and Perceptual Accuracy relative to Alpha Power

To ensure that the effects observed in relation to IAF were not confounded by oscillatory power, we conducted several analyses (See Supplementary Materials). First, we extracted time-frequency oscillatory amplitude maps for each participant from the specific electrodes used to compute IAF. Time-frequency decompositions were performed for 50 linearly spaced frequencies ranging from 2 Hz to 50 Hz. Wavelets with 3 cycles at the lowest frequency and 11 cycles at the highest frequency were applied. After obtaining the absolute value of the resultant analytic signal, we conducted a cluster-based permutation test to evaluate whether oscillatory amplitude varied based on response accuracy (correct vs. incorrect trials). Next, we extracted the trial-by-trial amplitude from the time-frequency analysis, averaging it within the -800 ms to -100 ms window, at the frequency closest to each participant's previously determined IAF. We performed regression analyses using a least-squares approach, incorporating both trial-by-trial fluctuation in IAF and alpha power as predictors of perceptual accuracy. This analysis allowed us to determine whether the influence of IAF on performance was independent of alpha power. We also replicated a bin-based analysis by dividing trials into terciles based on the previously extracted trial-by-trial alpha amplitude. We compared signal detection theory metrics across the lowest and highest terciles to confirm that the effects observed with IAF did not hold when considering variations in alpha power. Lastly, we conducted a time-frequency analysis using the same parameters as those applied in the comparison of correct and incorrect responses. This analysis aimed to assess potential differences in oscillatory amplitude between participants categorized into low and high IAF groups.

Here we include the new part and figure included in the supplementary materials:

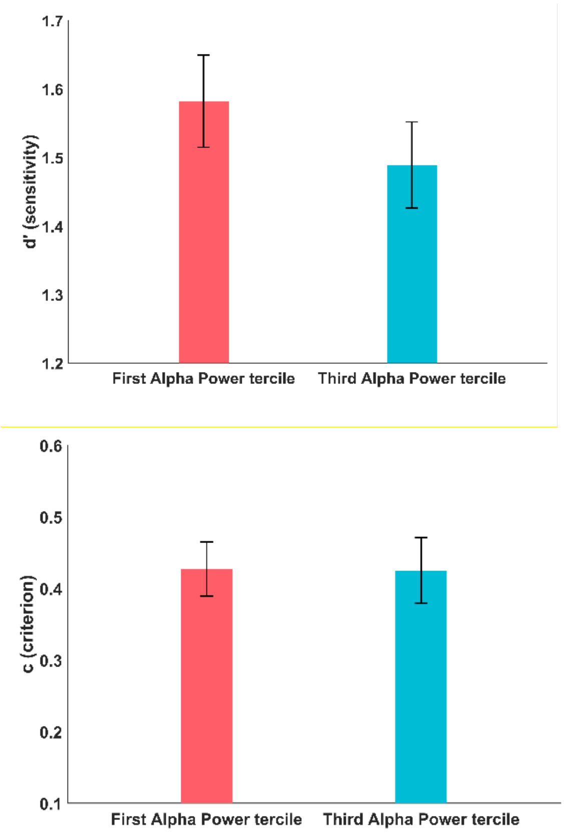


Figure S1. No significant differences observed in behavioral performance between first and third power terciles
We replicated the initial analysis included in the main text by binning trials based on pre-stimulus power in the first and third terciles. The analysis revealed that this binning approach did not dictate behavioural performance (all $t_{115} < 1.68$, all $p > 0.09$, all BF < 0.41).

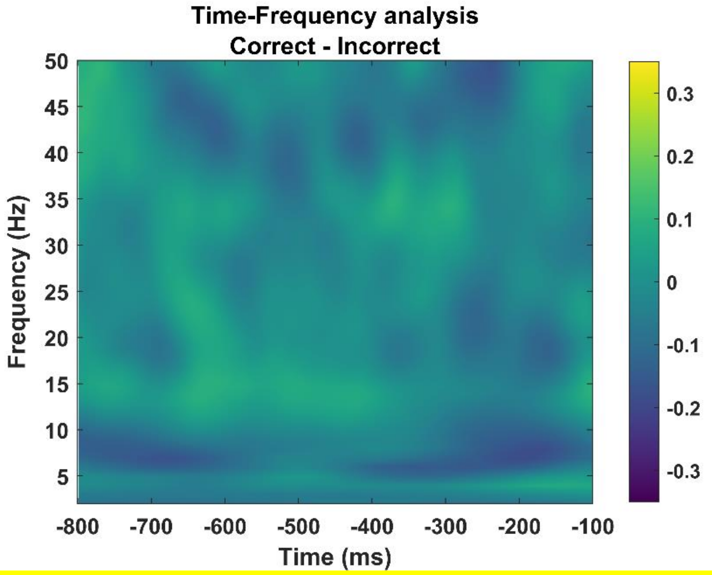


Figure S2. Correct and Incorrect trials are not characterized by difference in oscillatory amplitude.

We conducted a time-frequency analysis to assess whether correct responses exhibited different amplitudes compared to incorrect responses. The cluster-based analysis revealed no significant differences between the two types of trials, indicating that power, in contrast to IAF, is not capable of differentiating correctness in the trials.

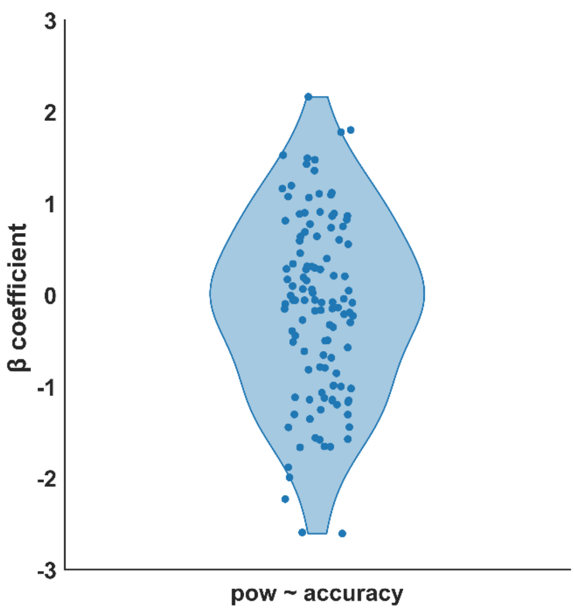


Figure S3. Trial-by-trial fluctuations in alpha power do not account for perceptual accuracy.

We conducted an analysis to evaluate whether trial-by-trial fluctuations in power, alongside those of IAF, influenced the accuracy of the responses provided. We extracted the power calculated from the time-frequency analysis by taking the alpha frequency closest to the IAF, normalizing it through a z-score transformation. Subsequently, we created a design matrix that included the intercept, IAF, and normalized alpha power. The regression analysis confirmed the significant predictive ability of IAF fluctuations on perceptual accuracy (mean slope = 0.37 $p < 0.001$). In contrast, fluctuations in power did not modulate the participant's ability to accurately retain the stimulus (mean slope = -0.15 p = 0.12).

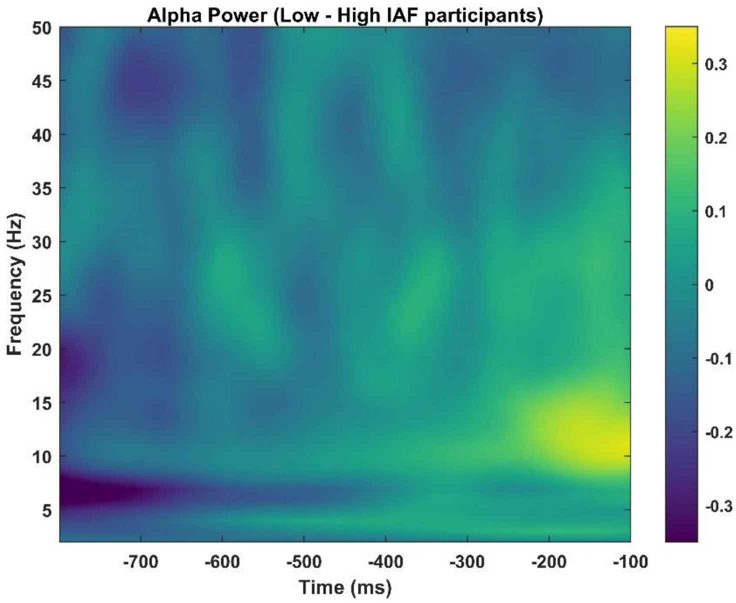
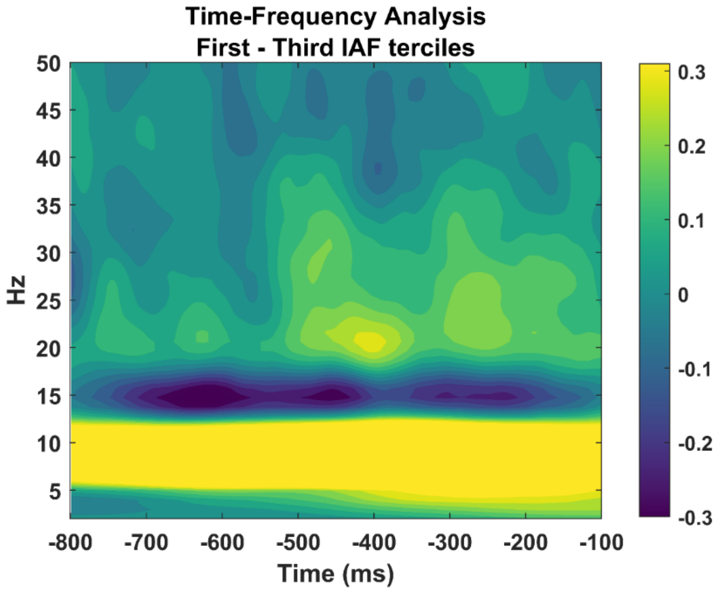


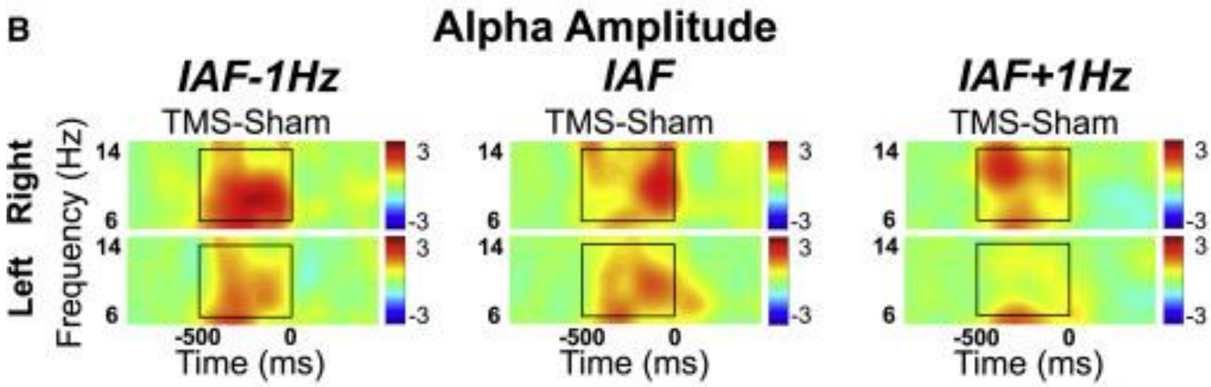
Figure S4. Low and Fast IAF groups do not exhibit significant differences in oscillatory amplitude.

We conducted a time-frequency analysis to assess whether participants in the low IAF frequency group exhibited different amplitude profiles compared to those in the high IAF frequency group. The cluster-based analysis revealed no significant differences between the two populations, indicating that power does not effectively differentiate between the two groups.

Finally, we appreciate the reviewers' insightful comments regarding the potential presence of a power difference between trials included in the 1st and 3rd alpha bins by implementing a time-frequency analysis. In the time-frequency map presented below, the difference between the TF maps extracted from the first bin versus the third bin is shown. It is relevant to note that the map has two polarities. Specifically, trials included in the first bin show higher power in the lower frequency alpha bands, while trials in the third tercile demonstrate an increase in power within the high alpha-low beta range.



We believe that these power differences are likely due to the two types of trials being characterized by different peak frequencies (e.g., higher in the third tercile), which subsequently could impact power (e.g., more power in the high alpha band). To illustrate this, we include a figure from a previous publication by our group (Di Gregorio et al., Curr. Biol.,2022), where we manipulated alpha frequency using rhythmic transcranial magnetic stimulation (rTMS) to increase or decrease it. As shown in the TF maps (the ones deriving from the analysis of right sensor i.e., the side on which the rTMS was applied), in the IAF + 1 Hz condition, there is a tendency for power to cluster at higher frequencies, whereas the opposite effect is observed in the IAF – 1 Hz condition.



In any case, although differences in power emerged when comparing the low and high alpha bands in the first and third terciles, these power variations do not account for the observed relationship between IAF and perceptual sensitivity. As discussed earlier, all control analyses, where power was included as a predictor of accuracy or sensitivity, failed to show any significant associations. Moreover, when power was incorporated as a factor in the trial-by-trial analysis, the relationship with IAF remained robust, retaining the same predictive strength as indicated by the Bayes factors, while power showed no meaningful relationship with accuracy. Given these findings, we are confident that power alone does not predict decision accuracy or perceptual sensitivity.

9. The authors should at least mention which sensors were used in the analyses at the beginning of the results section. I understand that the details of sensor selection belong to the methods but knowing which electrodes were used for the analyses is fundamental to interpret the results. Moreover, in the results section they describe how electrodes were selected but it would be informative to provide the final distribution of selected electrodes across participants (i.e. how many times was each electrode selected). Also, if I understand correctly, only one electrode was used for each participant? This could be more clearly stated.

Thank you for pointing out this aspect. We recognized the importance of spelling out the specific data analysis selecting the electrodes. To this end, we have added in the results section that:

> *We focused our analysis on a pool of occipito-parietal electrodes located on the midline and right side (i.e., Oz, POz, O2, PO4, PO8), selecting, for each subject, the electrode that exhibited the maximum alpha power in the prestimulus window. These electrodes were chosen because the stimulus was consistently presented on the left.*

Moreover, in the method section, we reported the final distribution of selected electrodes across participants:

> *The distribution of selected electrodes was as follows: Oz (11%), POz (25%), O2 (11%), PO4 (22%), and PO8 (31%).*

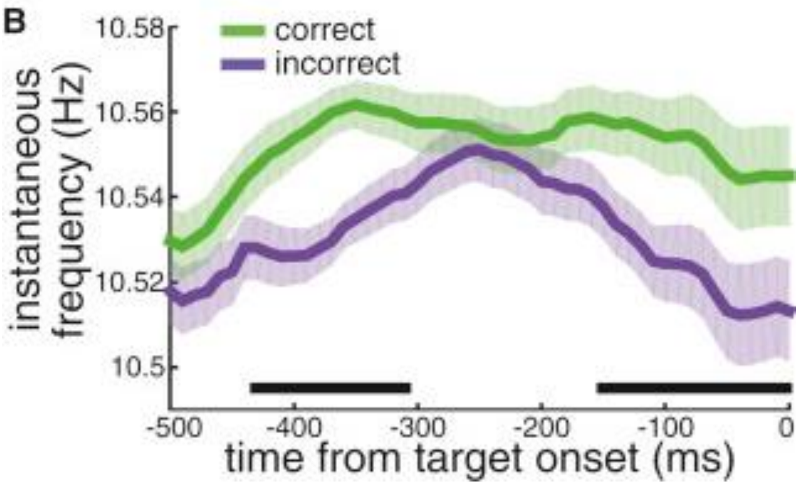
10. The authors do not discuss potential origins of individual differences in IAF. Although there might not be much to elaborate on in the literature (one potential factor I know of is white matter structure (Minami et al., 2020)), this is a central aspect of their study and I believe it would be interesting to at least propose future avenues for research on this question.
Minami, S., Oishi, H., Takemura, H. & Amano, K. Inter-individual differences in occipital alpha oscillations correlate with white matter tissue properties of the optic radiation. eNeuro 7, 2 (2020).

Thank you for your thoughtful comment, which raises an important aspect of our study. We agree that understanding the origins of individual differences in IAF is crucial, especially as these differences likely play a significant role in the IAF-phase relationships we have observed. Firstly, white matter structure is a key factor in IAF variability across subject, as indicated by the study of Minami et al. (2020) that you pointed out. Specifically, they found that the integrity of white matter tracts affects the frequency of oscillatory activity within the alpha band. This structural variability could therefore have a direct impact on how phase interact with IAF, potentially explaining some of the differences observed in our study. In addition to structural factors, genetic predispositions also appear to influence IAF (Smith et al., 2006) as well as an age-related decline in IAF has been consistently reported (Klimesch, 1999). This age-related variability might explain differences in how the IAF-phase relationship manifests across different age groups. Additionally, psychopathological conditions like schizophrenia (Ramsay et al., 2021), may also disrupt typical IAF dynamics, leading to altered IAF-phase interactions. Collectively, these factors are crucial to understanding the variability in IAF and how it interacts with phase to influence cognitive processes. In the revised manuscript, we have integrated this discussion to highlight the potential origins of IAF variability and its implications for the mechanisms we have identified:

> *Our findings underscore the importance of considering individual differences in IAF when interpreting phase-related effects. Inter-individual IAF variability is influenced by several factors, which could, in turn, modulate the IAF-phase relationship founded. For instance, the properties of the optic radiation—the primary white matter tract connecting the thalamus to the primary visual cortex—as well as genetic factors play a role in modulating individual differences in IAF (Minami et al., 2020; Smith et al., 2006). IAF also tends to decline with age (Klimesch, 1999), and conditions such as schizophrenia are associated with a lower IAF (Ramsay et al., 2021). Future research should explore how these IAF-related factors might modulate the role of phase in perceptul processes, either enhancing or diminishing its impact.*

11. The 3rd IAF bin in Fig1 A & B seems to fluctuate rhythmically and the two significant clusters in Fig. 1A are roughly separated by a theta cycle (300-250ms). Could the authors comment on this? More generally, the authors do not elaborate on potential cross-frequency coupling although numerous studies have linked theta-alpha coupling in visual perception. I do not believe that this is mandatory but it would be interesting to discuss these potential interactions.

Thank you for your insightful comment. This observation is indeed very interesting and aligns well with the rhythmic fluctuations seen in our data, where accuracy and sensitivity seem to oscillate with peaks and troughs every 250/300 ms (Figure 2). Notably, this rhythm echoes what was reported by Samaha et al., 2015 (see attached figure).



We had not initially considered cross-frequency interactions, but your suggestion raises an intriguing possibility. It seems plausible that theta-alpha cross-frequency coupling could modulate the influence of IAF on perceptual performance. For instance, if alpha is nested within the theta cycle, it could amplify the impact of IAF on perception, whereas a reduction in this coupling might dampen that effect. While investigating this in depth is beyond the scope of this study, we have added a note in the discussion, as we believe this is a very interesting direction for future research and could really shed light on the dynamics between these oscillations and perceptual processes.
Specifically, we have added in the discussion of the new version of the manuscript that:

> *Moreover, the interaction between IAF and phase is not the only factor that could influence the role of IAF in shaping perceptual accuracy. For instance, our data reveals rhythmic fluctuations in accuracy and sensitivity, with cycles occurring approximately every 250-300 ms (Figure 2A, 2B). Interestingly, this pattern mimics the one observed in Samaha et al. (2015) (see Figure 4), suggesting that cross-frequency coupling—particularly interactions between theta and alpha oscillations—might play a role in modulating the effect of IAF on perceptual performance (Canolty and Knight, 2010). While this study does not directly investigate these oscillatory dynamics, future research could explore how theta-alpha coupling influences the relationship between IAF and visual processing.*

12. It would be interesting to discuss how the present findings relate to computational models of (alpha) oscillations (e.g. Karvat & Landau, 2024, although this specific study is concerned with temporal integration).
Karvat, G., & Landau, A. N. (2024). A role for bottom–up alpha oscillations in temporal integration. Journal of Cognitive Neuroscience, 36(4), 632-639.

Thank you for this insightful suggestion, which we are pleased to integrate into our manuscript. The theories presented by Karvat and Landau (2024) resonate strongly with our findings. They highlight that in tasks requiring precise temporal resolution, the discriminative threshold between one and two flashes falls within the 15–40 ms range. This brief interval emphasizes the importance of phase, suggesting that effective temporal integration occurs not only when stimuli fall within the same cycle but potentially within one-eighth to one-fifth of the alpha cycle. Therefore, for stimuli to be significantly influenced by alpha oscillations, they must align with specific phases of the alpha wave. This aligns with our assertion that a higher IAF may enable stimuli to synchronize more closely with these critical phases, thereby enhancing perceptual performance in our detection task. Additionally, we have included considerations prompted by the seminal work of Haegens et al., 2014. Specifically, the authors observed that IAF gradually accelerates in response to increasing task demands, peaking from rest to low-demand conditions and reaching its maximum in high-demand scenarios. This pattern may indicate a mechanism through which the brain optimizes processing by increasing the likelihood of entering optimal phases for task performance.
We have incorporated these insights into the revised manuscript. Specifically, we have added the following points in the discussion:

> *Moreover, the relationship between IAF and phase aligns with recent theories proposed by Karvat and Landau* (2024)*. The authors noted that in tasks assessing temporal resolution, the discriminative threshold between one and two flashes falls within the range of 15–40 ms. This relatively short interval further supports the crucial role of phase. Specifically, if temporal integration occurs within a window corresponding to one-eighth to one-fifth of the alpha cycle, stimuli would need to reach specific parts of the alpha wave (i.e., phase) to be effectively influenced by*

> *alpha oscillations. In the context of our task, this suggests that higher IAF may enable stimuli to align more precisely with these critical wave phases, thus facilitating precise detection performance. Additionally, the gradual acceleration of IAF observed in response to increasing task demands—peaking as task demands rise from rest to low-demand conditions and reaching its maximum in high-demand scenarios* (Haegens et al., 2014)—*may reflect a mechanism that allows the brain to enhance the likelihood of re-entering optimal processing phases.*

13. In the "Instantaneous alpha phase shapes perceptual accuracy" section, no statistic or p-value is given. The authors should provide some information on the statistical results there.

Thank you for your comment. We have now reported all p-values derived from the cluster-based analyses, not only in the section highlighted by the reviewer but throughout the entire manuscript. Additionally, we have described in the Methods section how the p-values were calculated, providing details on the statistical approach employed for the cluster-based permutation tests.

> *We also calculated the p-value for each cluster by determining the position of the observed cluster within the distribution of clusters generated from the permutation procedure. This approach was applied consistently across all permutation-based analyses.*

14. There are a number of typos and errors across the manuscript (e.g. L.249 "invstgate", the citation of the DDM package used is Wiecki, not Wiekli, and it is not included in the references, etc.)

Thank you for pointing out these issues. We have corrected the typographical errors and updated the missing reference for the DDM package.

15. A perceptually homogenous color-scale should be used in Fig.4C (https://www.mathworks.com/matlabcentral/fileexchange/51986-perceptually-uniform-colormaps)

We have used a perceptually homogeneous color scale in all figures where a color code is applied. Specifically, we employed the *Viridis* color map to ensure uniformity.

16. The title of section "Alpha-Phase clustering increases following correct responses only in slow IAF individuals." is confusing as the effect that is shown is post-stimulus presentation but not necessarily post-response (or at least we cannot know since RTs are not reported).

Thank you for pointing out the potential confusion with the section title. We agree that the original title may have been misleading. We did not intend to imply that the effect occurs strictly after a correct response in a temporal sense. Rather, our intention was to convey that the observed alpha-phase clustering effect was present in trials where the response was correct.

To better reflect this nuance, we amend the title with the new one:

> *Alpha-phase clustering increases for correct responses only in low IAF individuals.*

Response to Reviewers

Reviewer #1:

The authors have made significant efforts to revise their manuscript. However, several methodological concerns raised previously remain valid and require further attention. Below, I briefly outline the key issues along with recommendations for improvement.

1. The main conclusions still rely on computing a so-called 'instantaneous alpha frequency' by applying the Hilbert transform to obtain phase and then computing the derivative. However, this approach is flawed. The term 'instantaneous frequency' is misleading in the sense that frequency can accurately be computed only over finite time windows. Instantaneous frequency estimates are not consistent with fundamental principles in physics, as greater frequency resolution requires longer time windows. Michael Cohen who initially proposed this method showed in subsequently published research that this approach leads to biased estimates that depend strongly on 1/f noise and SNR. While the authors attempted to remove 1/f noise, their approach appears to be problematic. As shown on p.12 in their rebuttal, the PSD after 1/f noise removal shows substantially lower alpha power (an order of magnitude lower), which is rather surprising given the original PSD. To ensure unbiased frequency estimates, it is crucial to adhere to established methods such as FFT over longer time windows. This approach is widely validated and avoids the pitfalls associated with 'instantaneous frequency' calculations.

We appreciate the reviewers' detailed comments and agree that it is useful to confirm that our findings linking alpha frequency with perceptual accuracy are robust and not based on a single method of extraction, i.e., the instantaneous frequency approach. To address these concerns, we implemented new set of analyses that confirmed the reliability of the finding derived from the instantaneous frequency calculation:

ALTERNATIVE ANALYSIS 1: As an initial complementary analysis, we pooled trials from each condition (correct versus incorrect) to circumvent the limitations of single-trial estimates, which can be vulnerable to noise due to the brief prestimulus period. This approach enables us to obtain a more stable, high signal-to-noise estimate of the alpha peak, thereby mitigating the potential imprecision inherent in single-trial analyses. Therefore, for each subject we visually inspected the pooled power spectra for correct and incorrect trials separately and only included those subjects for whom the alpha peak was clearly identifiable in both conditions (excluded participants N = 15 out of 116). t-tests comparing the extracted IAF values revealed that correct responses were associated with a significantly higher IAF, underscoring the relationship between elevated alpha frequency and task performance.

ALTERNATIVE ANALYSIS 2: Moreover, we extended our analysis by applying an automated, data-driven algorithm—following the approach of Corcoran et al. (2018)—to extract IAF from the same pooled power spectra. This approach first smooths the power spectrum using a Savitzky-Golay filter, which is critical for reducing random noise and revealing the underlying spectral structure. Following smoothing, the algorithm computes the first derivative of the spectrum with respect to frequency, identifying the local maximum in the alpha band as the IAF. This objective method reduces the subjectivity inherent in visual inspection. Importantly, the algorithm is designed to be highly conservative, providing an estimate only when the signal is robust, thereby preventing unreliable peaks from being forced into the analysis (N = 34 out of 116). Despite the strictness of this approach, the resulting estimates still revealed a significantly higher IAF in correct trials, confirming the robustness of our findings independently of the extraction method adopted. Notably, alternative analyses 3 and 4 (see below) using a trial-by-trial FFT-based estimation approach yielded the same results without requiring the exclusion of any participant.

ALTERNATIVE ANALYSIS 1 and 2 have been added in the new version of the supplementary materials:

> ***IAF extracted from the pooled correct trials was higher than that extracted from the pooled incorrect trials.***
>
> Beyond single-trial analyses, we also examined whether IAF differences between correct and incorrect trials were evident when estimated from pooled data. This approach provides a complementary perspective, as it avoids trial-by-trial variability and instead focuses on broader spectral properties of the prestimulus signal.

Specifically, we computed the power spectrum by pooling all trials within each response category (correct vs. incorrect) and identified the alpha peak separately for each condition. Our results revealed a systematic difference: IAF extracted from the pooled correct trials was significantly higher than that extracted from the pooled incorrect trials. This finding was consistent across subjects and was confirmed using both visual inspection of the power spectrum ($t_{100} = 3.11$, $p < 0.01$, BF = 9.80) of the spectra and the Corcoran et al. (2018) algorithm for automated peak detection applied to the aperiodic-free power spectrum ($t_{81} = 2.77$, $p < 0.01$, BF = 3.73). Importantly, this pattern replicated the results obtained using the single-trial approach, reinforcing the conclusion that prestimulus IAF is a key predictor of perceptual accuracy. The consistency across different estimation methods highlights the robustness of the observed effect, ruling out the possibility that it arises from noise-related fluctuations in single-trial analyses.

ALTERNATIVE ANALYSIS 3. Secondly, also following the insights from reviewer 4, we verified the single-trial IAF estimates using complementary FFT-based approaches as included in the new version of the methods section:

Additionally, we adopted complementary approaches to examine the role of IAF without relying on the instantaneous frequency method, instead using FFT-based frequency estimation. First, we computed the power spectrum from the pre-stimulus signal (-800 to -100 ms), matching the time window used for the instantaneous frequency approach. We then applied Corcoran's algorithm to the single-trial power spectra to extract the single-trial IAF. This peak detection method, performed directly on the frequency spectrum, helps overcome potential limitations of the instantaneous frequency approach (Samaha and Cohen, 2022; Schoffelen et al., 2024).

Specifically, we adapted the Corcoran algorithm and applied it separately to each single trial. Subsequentially, we replicated both the tercile-based and trial-by-trial analyses using these single-trial IAF estimates. For the tercile-based analysis, trials were divided into three groups based on the prestimulus IAF, and performance metrics were compared between the lowest and highest terciles. Consistent with our previous results, while no effect was observed on criterion, both sensitivity and accuracy were significantly higher in the third IAF tercile. In the trial-by-trial analysis, fluctuations in IAF on individual trials were used to predict perceptual performance, and the results confirmed that higher IAF on a given trial corresponded to an increased probability of correctly detecting the presence/absence of the target.

These new analyses have been added in the new version of the supplementary materials:

FFT-based method for estimating single-trial IAF confirmed the pattern of results observed with the instantaneous frequency approach.

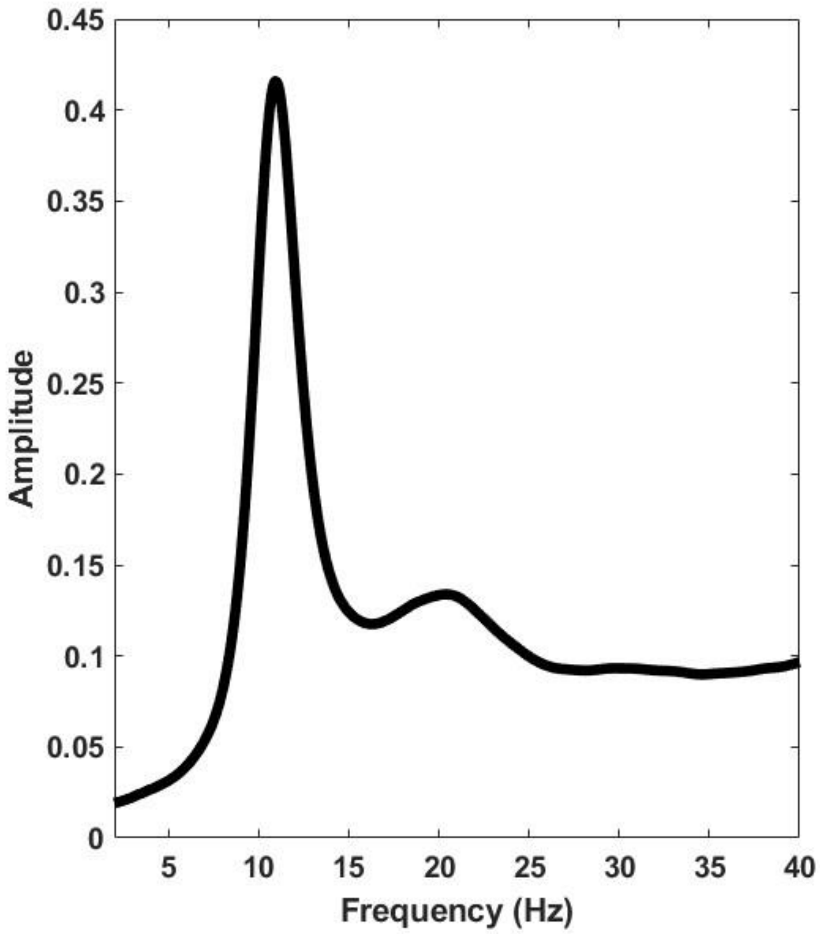
To further validate the reliability of single-trial IAF estimation, we compared the results obtained using the FFT-based method with those derived from the instantaneous frequency approach. Specifically, we applied the Corcoran et al. (2018) algorithm to the single-trial power spectra, allowing us to extract IAF values based on a robust peak detection method rather than phase-based instantaneous frequency calculations. The key advantage of this approach is its reliance on spectral features rather than time-domain fluctuations, reducing the influence of transient noise. Our analyses revealed a strong correspondence between the two methods. Specifically, trials in which the FFT-based IAF estimates were in the third terciles showed higher accuracy ($t_{115} = 3.03$, $p < 0.01$, BF = 7.64) and sensitivity ($t_{115} = 2.77$, $p < 0.01$, BF = 3.85), while maintaining the same level of criterion ($t_{115} = 1.46$, $p = 0.15$, BF = 0.29). Moreover, fluctuation in trial-by-trial FFT-bases IAF estimate impacted the probability of making a correct detection of the target (t_{115} =, $p < 0.01$, BF = 40.31). Therefore, both methods consistently indicated that higher IAF was associated with better perceptual performance, confirming the robustness of our original findings. These results suggest that different analytical frameworks can reliably capture the same underlying neural dynamics, reinforcing the conclusion that prestimulus IAF fluctuations play a crucial role in shaping perceptual outcomes.

ALTERNATIVE ANALYSIS 4. Following the request of reviewer 4, we further extended our analysis by implementing a Gaussian function-based approach to estimate the IAF on a trial-by-trial basis. In this method, a Gaussian curve is fitted to the single-trial power spectrum. The peak of this fitted Gaussian is taken as the IAF for that trial. Building on these single-trial estimates, we conducted complementary analyses. In the tercile-based analysis, trials were sorted into three groups based on their prestimulus IAF values. Performance metrics were then compared between the lowest and highest terciles; consistent with our previous results, while criterion remained unaffected, both sensitivity and accuracy were significantly higher in the highest IAF tercile. In the trial-by-trial analysis, we examined how fluctuations in IAF predicted perceptual performance. The results confirmed that a higher IAF on any given trial corresponded to an increased probability of correctly detecting the presence or absence of the target. Moreover, correct trials were associated with higher IAF compared to incorrect trials. These new analyses have been added in the new version of the supplementary materials:

Single-Trial IAF Estimation via Aperiodic-Free Gaussian Fitting

To further validate the reliability of our single-trial IAF estimation, we conducted an additional analysis using an aperiodic-free Gaussian fitting approach. In this procedure, the power spectrum was computed over the prestimulus window using Welch's method per each trial separately, focusing on the alpha band. An initial peak estimate was determined by identifying the frequency with the highest power within this band. Around this peak, we defined a narrow fitting window extending 3 Hz on either side to isolate the region of interest. Within this window, a Gaussian function—characterized by its amplitude, centre frequency, and width—was fitted to the data, and the centre frequency of the fitted Gaussian was taken as the IAF for that trial. An adjusted R^2 was calculated to assess the goodness-of-fit, and trials were excluded if the adjusted R^2 was below 0.9 or if the IAF was estimated at the boundaries of the fitting window, suggesting that the fit was constrained rather than reflecting a true spectral peak. First, we also compared the goodness-of-fit between correct and incorrect trials and observed no significant differences (Adjuster r^2 correct = 0.98± 0.01; Adjuster r^2 incorrect = 0.98 ± 0.01; t_{115} = -0.33, p = 0.74). Moreover, the IAF estimates obtained with this method successfully predicted participants' performance, replicating the results presented in the main text. Specifically, terciles analysis confirmed that trials included in the third vs. first terciles are characterized by higher sensitivity and accuracy (all t_{115} > 2.96, all p < 0.01; all BF > 6.46). Moreover, correct responses were associated with higher IAF compared to incorrect responses (t_{115} = 2.65, p < 0.01; BF = 2.91) and trial-by-trial fluctuations in IAF predicts variations in perceptual accuracy (t_{115} = 2.69, p < 0.01; BF = 3.24)

Alternative analysis 5. Regarding the comment relative to the appropriateness of removal of 1/f component, we would like to emphasize that, although the absolute alpha power appears reduced following 1/f noise removal, this decrease does not reflect a lower signal-to-noise ratio (SNR). We computed the relative alpha power, which compares the alpha power to the total power across the spectrum. This ratio inherently offers a more precise indices of the alpha signal's prominence relative to the background. Our analysis revealed that after removing the 1/f component, the relative alpha power increased (38%) compared to the spectra where the 1/f component was retained (33%), indicating that the alpha signal is more clearly distinguished from the background noise. Moreover, we employed an alternative approach to remove 1/f, by applying a polynomial fitting procedure to the power spectrum. First, for each trial, we computed the FFT to obtain both the power spectrum and the phase information, then converted the power spectrum into a logarithmic scale. In this log-transformed space, we fit a second-degree polynomial to capture the 1/f component. By subtracting this polynomial from the log power spectrum, we effectively removed the 1/f noise (see figure below).



Finally, we recombined the corrected power spectrum with the original phase information and performed an inverse FFT to reconstruct the time-domain signal without the 1/f contribution. We computed the relative alpha power and, even using this alternative approach, it is superior compared to the original signal (40%). Moreover, we replicated the original analysis linking IAF to perceptual precision when extracting IAF from this aperiodic-free signal, as detailed in the supplementary materials:

Confirmation of the Robust IAF–Sensory Precision Relationship Using an Alternative 1/f Removal Method

We conducted supplementary investigations to assess whether the results linking IAF and sensory accuracy are robust to different methods for removing the 1/f component. Specifically, we applied an alternative approach that uses a polynomial fitting procedure to remove 1/f noise. For each trial, we computed the FFT using Welch's method to obtain both the power spectrum and the phase information over the prestimulus window. The power spectrum was then converted into a logarithmic scale, and a second-degree polynomial was fitted to capture the 1/f component. By subtracting this polynomial from the log-transformed power spectrum and converting the result back to the original scale, we effectively removed the 1/f noise. The corrected power spectrum was then recombined with the original phase information and transformed back into the time domain using an inverse FFT.

We subsequently replicated our original analysis linking IAF to perceptual precision using the IAF estimates derived from this aperiodic-free signal. The results confirm that our findings are robust and consistent regardless of the specific method used for 1/f removal. Specifically, tercile analysis confirmed that trials in the third versus the first tercile are characterized by higher sensitivity and accuracy (all $t_{115} > 2.80$, all $p < 0.01$; all BF > 4.16). Moreover, correct responses were associated with higher IAF compared to incorrect responses ($t_{115} = 4.29$, $p < 0.01$; BF = 439.41), and trial-by-trial fluctuations in IAF predicted variations in perceptual accuracy ($t_{115} = 4.42$, $p < 0.01$; BF = 703.19).

Collectively, these multiple lines of evidence demonstrate that the relationship between IAF and perceptual accuracy is robust and not an artifact of the instantaneous frequency method. By validating our initial results using both classical and data-driven spectral analyses, ensuring the reliability of single-trial estimates, and replicating our findings with independent analytical frameworks (FFT), we are confident that these additional analyses convincingly establish the validity of our original conclusions and underscore the genuine neurophysiological basis of the observed effects.

2. The authors suggest that despite the minute differences in IAF within each participant they observe substantial variability when computing variance pooled over time, trials and participants (p.2 in rebuttal). However, this approach conflates within trial variability across time, within-participant variability across trials and between-subject variability. Because the statistical inference (i.e. comparison: correct vs incorrect) is across trials within each participant, this conflation of variability is not appropriate.

3. In support of their own findings, the authors cite previous studies in high impact journals that reported comparable differences in alpha frequency. However, these older studies relied on the instantaneous alpha frequency estimation approach which has since then been shown to be flawed.

We thank the reviewer for these two comments that highlight the potential issue of conflating variability sources in our analysis. Although pooling variance over time, trials, and participants could mix within-trial, within-subject, and between-subject variability, our primary statistical inference is based exclusively on paired comparisons within each participant. This design ensures that our results reflect true within-subject effects, rather than an artifact of mixed variability. Importantly, with regard to the magnitude of the effect linking IAF and perception, evidence from electrocorticography (ECoG) studies (Chen et al., PLOS Biology, 2019)—which are less susceptible to spatial mixing and volume conduction than extracranial EEG—has demonstrated larger IAF-related effects (up to 0.20 Hz) in perceptual decision tasks. In contrast, the EEG analysis in the same study revealed a significant, yet weaker, effect (0.06 Hz) when compared to that reported in our paper. This finding suggests that the effect we observed at the EEG level might be a conservative estimate due to inherent signal mixing, yet it still robustly reflects an underlying neurophysiological mechanism.

Moreover, in response to concerns regarding the use of the instantaneous alpha frequency approach, we have complemented our analysis with more classical, FFT-based approach (see point 1). The persistence of the observed differences in alpha frequency between conditions—even when using established method for estimating the peak—attests to the robustness of our results. Moreover, just to point out, several recent studies (e.g., Han et al., Cerebral Cortex, 2023; Morrow et al., Biorvx, 2024; Santoni et al., NeuroImage: Clinical, 2025; Buergers and Noppeney, Nature Human Behaviour, 2022) continue to employ instantaneous frequency measures, and they consistently report convergent results when cross-validated with FFT-based methods (Menétrey et al., Imaging Neuroscience, 2024). Our own trial-by-trial analysis further supports this, as we found a strong correlation between instantaneous frequency and FFT-derived trial-by-trial IAF estimates. This further analysis has been added in the supplementary materials:

The FFT-based and instantaneous frequency approaches result in comparable single-trial IAF values.

A direct comparison of single-trial IAF values obtained through the FFT-based and instantaneous frequency methods revealed a significant association between the two approaches ($r = 0.65$, $p < 0.01$, $BF > 1000$). This strong correspondence suggests that the two methods capture the same neural signal properties, albeit from different analytical perspectives.

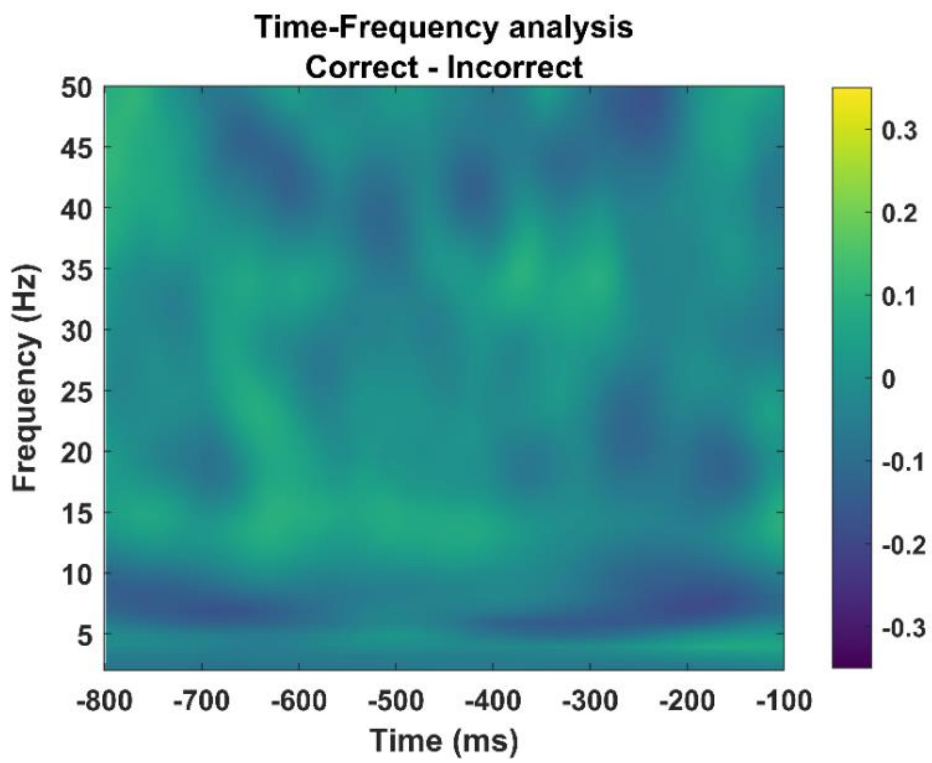
Therefore, the instantaneous frequency method yields estimates that are both significant and robustly associated with classical FFT-based approaches.

Lastly, the time-resolved nature of the instantaneous frequency approach allowed us to capture dynamic effects that standard FFT analysis might overlook. In particular, the temporal profile of the IAF-dependent effect (Figure 2A) closely mirrors that reported by Samaha et al. (2015), hinting at a potential theta–alpha cross-frequency interaction as insightfully suggested by reviewer 3 in the first round of revision. This insight, now incorporated into our discussion as a promising direction for future research, would likely have remained undetected using conventional FFT methods.

In summary, the convergence between the different analytical approaches not only verifies that our findings are not driven by a potentially artefactual procedure in alpha peak extraction but also reinforces the overall robustness and reliability of our results. We appreciate the reviewer's suggestion to perform an FFT-based analysis, which has provided an additional layer of verification to our conclusions.

4. Alpha power is another important confound. The authors show that there is not a significant difference in alpha power using classical statistics. However, there is non-decisive evidence using Bayesian analysis (in fact, even weak evidence for a power difference). Hence, alpha power confounds cannot be completely excluded.

Thank you for the comment. In response to your concern, also following suggestions from reviewer 4, we conducted additional analyses using a straightforward power spectrum analysis (see Figure S2) to further rule out any power-related effects. Specifically, we computed and plotted the power spectrum from 4–20 Hz and compared the average alpha power between correct and incorrect decisions, finding that alpha power did not differ significantly between the two trial types. This result reinforces the findings of the control analyses already included in the previous version, which demonstrated that trials with low vs. high alpha power were not associated with differences in perceptual accuracy. Indeed, the Bayes Factors were below 1, indicating evidence in favor of the null hypothesis. Furthermore, when directly comparing the IAF-related effect to the power-related effect, the Bayes Factor was, on average, at least 20 times larger for the IAF analysis. This makes it highly unlikely that such a negligible power-related effect could account for the strong and consistent IAF-related effects on perceptual accuracy observed in our study. We have included this new analysis and the new figure in the supplementary materials



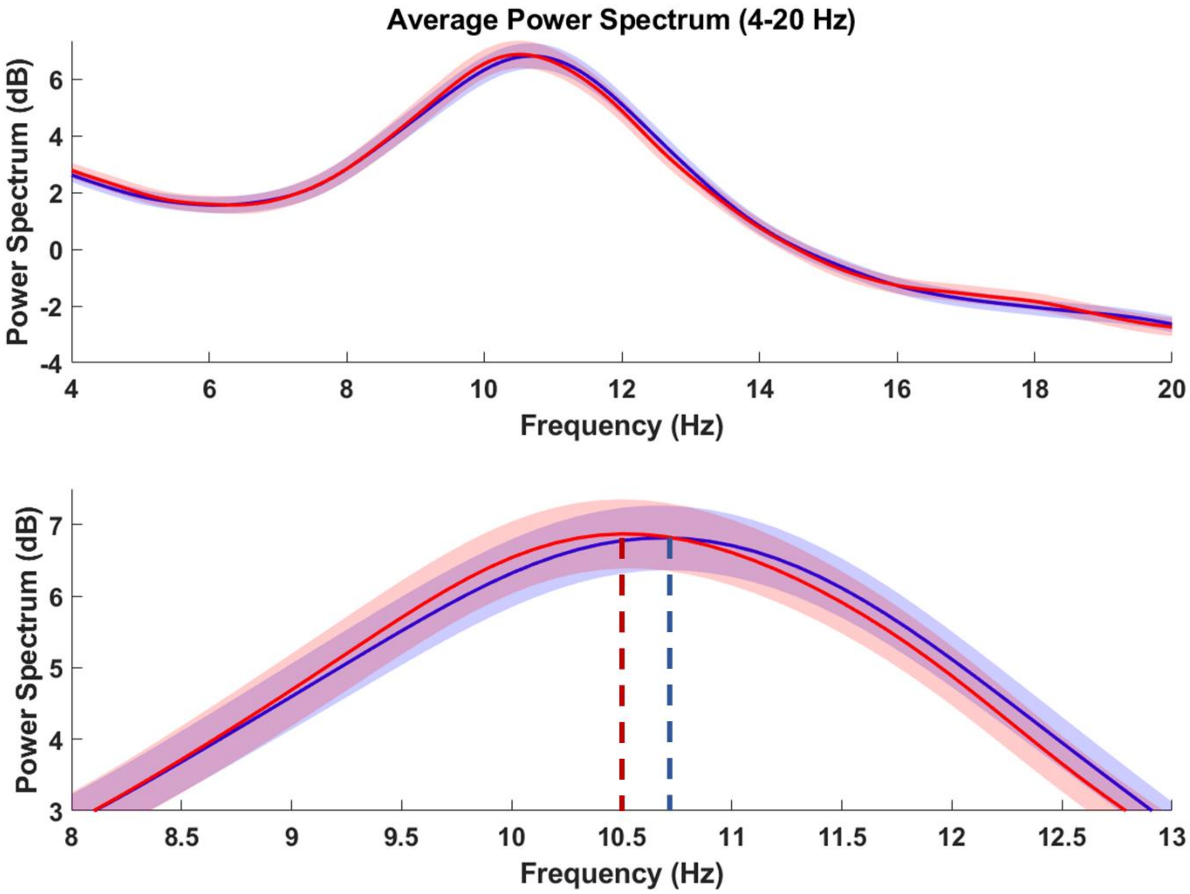


Figure S2. Correct and Incorrect trials are not characterized by difference in oscillatory amplitude.

We conducted a time-frequency analysis to assess whether correct responses exhibited different amplitudes compared to incorrect responses. The cluster-based analysis revealed no significant differences between the two types of trials, indicating that power, unlike IAF, does not differentiate correctness across trials. Additionally, we employed a more direct approach by computing the power spectrum for pooled correct vs. pooled incorrect trials. This allowed us to assess power spectrum differences between these two conditions within the 4-20 Hz range. In line with previous analysis, we found again no significant differences in power, neither in the alpha band nor in other frequency bands, including theta and beta (all t_{115} < 0.61, all p > 0.54, all BF < 0.12). Please also note the mean frequency difference between pooled corrected and incorrected responses of the order of magnitude similar to the one observed in the trial-by-trial analysis (~0.20 Hz difference between correct and incorrect responses).

Despite these concerns, the study may contribute to the field by sparking further research into the role of alpha frequency. However, ideally (and as previously suggested), the authors should address the methodological concerns and employ techniques known to provide unbiased alpha frequency estimates (i.e. FFT over finite timewindows).

We thank the reviewer for this comment. As noted in our previous responses, we have addressed these methodological concerns by incorporating alternative analyses—including FFT over finite time windows—that reinforce our original findings. We appreciate that this work can spark further research into the role of alpha frequency, and we are grateful for the reviewer's insightful comments across both rounds of review, which have helped us refine our research question and improve our methodology.

Reviewer #3:

The authors have addressed all my comments and suggestions in great details. I commend the authors on the work and care invested in responding to each comment or concern I had.

I believe the manuscript was greatly improved and that this is a well-designed and very impactful study.

We thank the reviewer, and we are delighted that you found our study both well-designed and impactful. Thanks again for the constructive comments that have significantly improved our work.

Reviewer #4:

I was invited to substitute for Reviewer 2 and assess whether the concerns raised have been appropriately addressed. First and foremost, I share the highly positive impression of the manuscript and would like to congratulate the authors on their outstanding work. The study represents an ambitious effort to scrutinize the relationship between IAF and visual perception, featuring a large sample size, multiple analytical approaches that converge on a main conclusion, and a conceptual integration of distinct research strands (perceptual and temporal sampling). Moreover, it generates testable research questions that will undoubtedly inspire future studies. Taken together, these strengths make the manuscript a strong candidate for publication in Nature Communications, which I strongly endorse.

A significant portion of Reviewer 2's comments focused on wording, typos, overlooked literature, and the need for clarifying remarks (e.g., regarding the strength of effects and the scope of the IAF-perception relationship, i.e., restriction to only very short stimuli). I believe the authors have done an excellent job addressing these concerns both in their response and in the revised manuscript.

Thank you for your positive and encouraging feedback. We appreciate your recognition of our efforts and are grateful for your support.

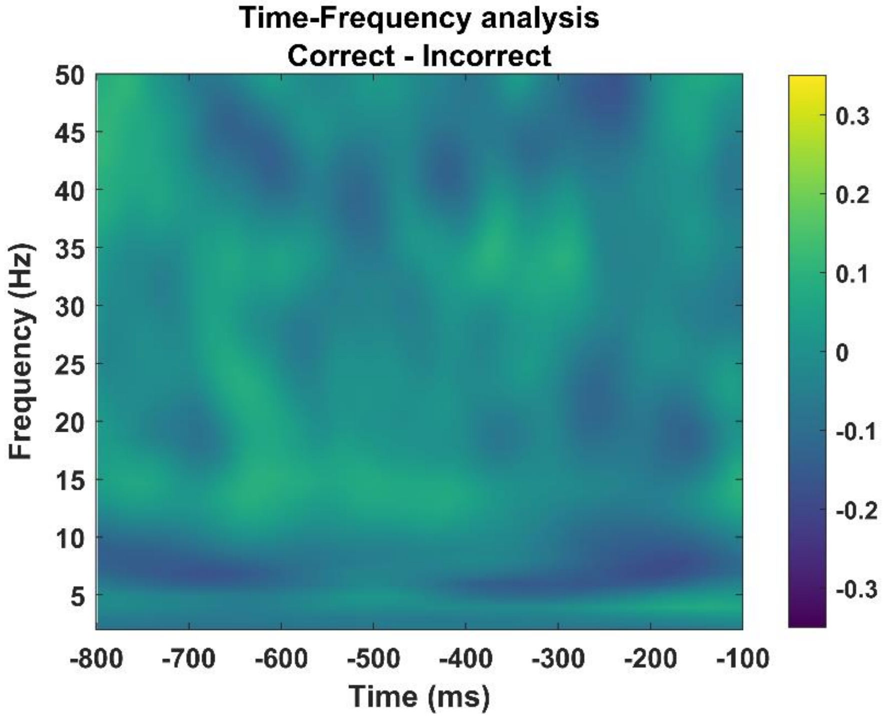
The most critical issue raised by Reviewer 2 concerns the potential influence of alpha power on instantaneous frequency. The authors employ multiple approaches to rule out a statistically significant impact on the IAF effect, though one could argue that the statistical approach in Figure S2 is overly conservative, as it tests a very broad frequency and time range.

However, my concern lies more with the reliability of extracting IAF at the single-trial level, which is crucial for several of the presented analyses. The methods section states that individuals without a clear IAF were excluded (ll. 582–584), but this does not fully account for various ways in which the power spectrum might influence the findings. For instance, changes in non-alpha frequency bands associated with more accurate perception could "leak" into the narrow-band filtered data, thereby affecting the estimated peak frequencies. To address this, it would be helpful to present Figure S2 in power spectral form over a narrower range (e.g., up to 20 Hz) for both conditions, overlaid with an estimate of variability.

Another potential issue is that in trials where participants are less accurate, noisier data could make alpha peak estimation less reliable. One way to test this would be perhaps to use the "aperiodic-free" single-trial power spectra and fit Gaussians to estimate peak frequency and Goodness-of-Fit per trial. If the effect is genuinely driven by IAF, Goodness-of-Fit should not differ between conditions.

Thank you for your insightful comments. In response, we have conducted a series of additional analyses and methodological checks to address your concerns regarding the potential influence of alpha power on the IAF effect and the reliability of single-trial IAF estimation. Below is a detailed explanation of our approach and findings.

A) To mitigate concerns that changes in non-alpha frequency bands could “leak” into our narrow-band filtered data and bias the estimated IAF, we re-examined the power spectra over a restricted frequency range as you suggested (up to 20 Hz). In the updated Figure S2, we overlay the power spectra for both correct and incorrect trials with an estimate of variability. The results reveal no significant differences in overall power—not only in the alpha band but also in other bands—between the two conditions. This analysis strengthens our claim that the differences observed in the IAF are not an artifact of spectral leakage from non-alpha frequencies.



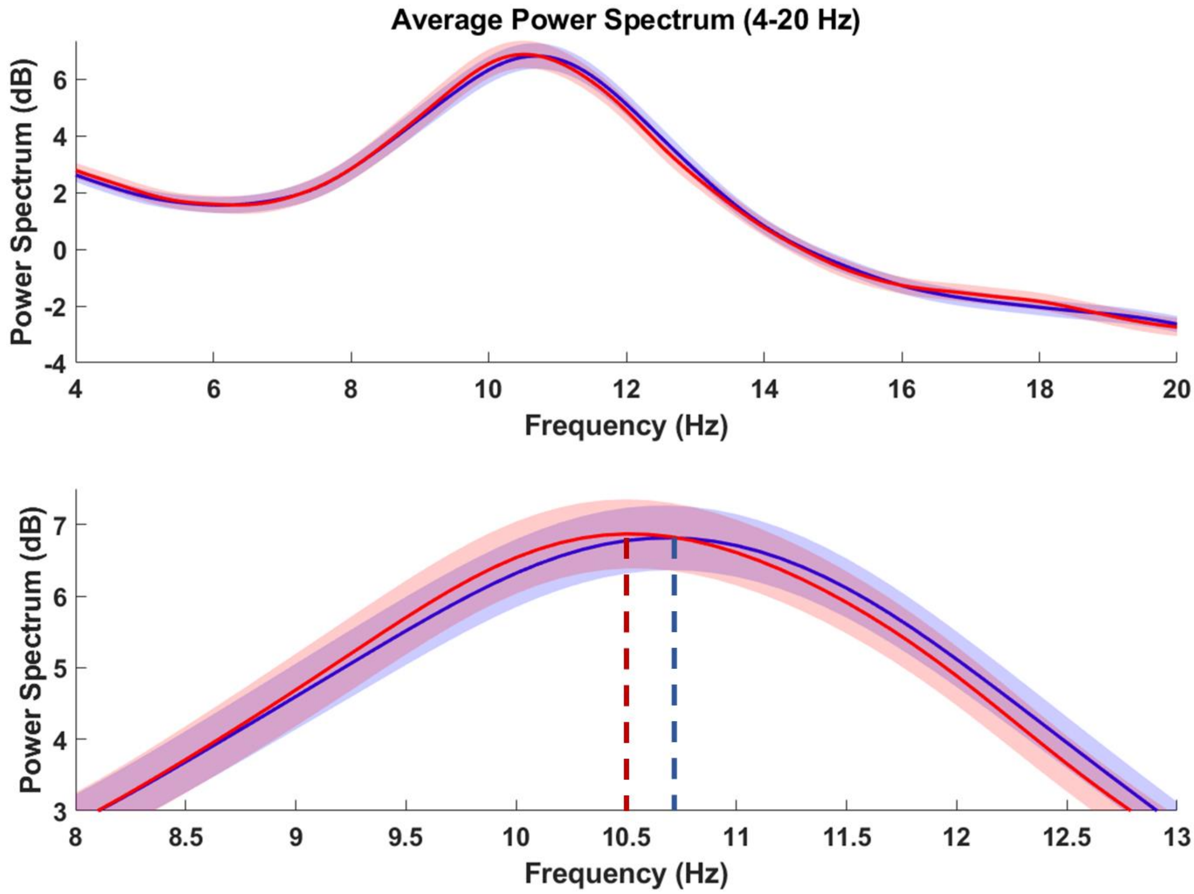


Figure S2. Correct and Incorrect trials are not characterized by difference in oscillatory amplitude.

We conducted a time-frequency analysis to assess whether correct responses exhibited different amplitudes compared to incorrect responses. The cluster-based analysis revealed no significant differences between the two types of trials, indicating that power, unlike IAF, does not differentiate correctness across trials.

Additionally, we employed a more direct approach by computing the power spectrum for pooled correct vs. pooled incorrect trials. This allowed us to assess power spectrum differences between these two conditions within the 4-20 Hz range. In line with previous analysis, we found again no significant differences in power, neither in the alpha band nor in other frequency bands, including theta and beta (all $t_{115} < 0.61$, all $p > 0.54$, all BF < 0.12). Please also note the mean frequency difference between pooled corrected and incorrected responses of the order of magnitude similar to the one observed in the trial-by-trial analysis (~0.20 Hz difference between correct and incorrect responses).

B) Given that single-trial IAF estimation is critical for several analyses in our study, we took additional steps to confirm its reliability. First, as proposed by the reviewer, we ran an aperiodic-free Gaussian fitting at the single-trial level. First, we computed the FFT using Welch's method to obtain the power spectrum over the prestimulus window. We then focused on the alpha band and identified the frequency with the highest power as an initial estimate of the peak. Around this initial peak, we defined a narrower fitting window—extending a 3Hz on either side—to isolate the region of interest. Within this window, we fitted a Gaussian function characterized by its amplitude, centre frequency, and width. The centre frequency from the fitted Gaussian was taken as the individual alpha frequency (IAF) for that trial, and we also computed an adjusted R^2 to assess the goodness-of-fit. To ensure the reliability of our IAF estimates, we excluded trials where the adjusted R^2 was below 0.9 and those in which the IAF was estimated at the boundaries of the predefined frequency window when the Gaussian fit may have been constrained by the limits of the fitting window rather than capturing a true spectral peak. Then, we investigated whether the goodness of fit differed between correct and incorrect trials. We found no significant differences in the fit quality, both when considering all trials and when using only those retained after applying our exclusion criteria. Moreover, the IAF estimates derived from this method successfully predicted participants' performance, replicating the results outlined in the initial version of the paper, as detailed in the supplementary material:

Single-Trial IAF Estimation via Aperiodic-Free Gaussian Fitting

To further validate the reliability of our single-trial IAF estimation, we conducted an additional analysis using an aperiodic-free Gaussian fitting approach. In this procedure, the power spectrum was computed over the prestimulus window using Welch's method per each trial separately, focusing on the alpha band. An initial peak estimate was determined by identifying the frequency with the highest power within this band. Around this peak, we defined a narrow fitting window extending 3 Hz on either side to isolate the region of interest. Within this window, a Gaussian function—characterized by its amplitude, centre frequency, and width—was fitted to the data, and the centre frequency of the fitted Gaussian was taken as the IAF for that trial. An adjusted R^2 was calculated to assess the goodness-of-fit, and trials were excluded if the adjusted R^2 was below 0.9 or if the IAF was estimated at the boundaries of the fitting window, suggesting that the fit was constrained rather than reflecting a true spectral peak. First, we also compared the goodness-of-fit between correct and incorrect trials and observed no significant differences (Adjuster r^2 correct = 0.98± 0.01; Adjuster r^2 incorrect = 0.98 ± 0.01; t_{115}

= -0.33, p = 0.74). Moreover, the IAF estimates obtained with this method successfully predicted participants' performance, replicating the results presented in the main text. Specifically, terciles analysis confirmed that trials included in the third vs. first terciles are characterized by higher sensitivity and accuracy (all t_{115} > 2.96, all p < 0.01; all BF > 6.46). Moreover, correct responses were associated with higher IAF compared to incorrect responses (t_{115} = 2.65, p < 0.01; BF = 2.91) and trial-by-trial fluctuations in IAF predicts variations in perceptual accuracy (t_{115} = 2.69, p < 0.01; BF = 3.24)

C) Moreover, we employed the algorithm proposed by Corcoran et al. (2018), which offers enhanced robustness and precision in detecting the alpha peak compared to traditional methods. The key features of this algorithm are that it begins by applying a Savitzky-Golay filter to smooth the estimated power spectrum. This smoothing process is essential because it reduces random noise and transient artifacts, revealing the underlying spectral structure more clearly. After smoothing, the algorithm calculates the first derivative of the spectrum. It then identifies the frequency at which the derivative changes from positive to negative—a transition that indicates a local maximum within the alpha band. This point is taken as the IAF. A critical advantage of this method is that it only returns an IAF estimate if the signal is sufficiently reliable. In cases where the smoothed spectrum does not exhibit a clear alpha peak (due, for example, to high noise levels or weak oscillatory activity), the algorithm refrains from forcing an estimation. This safeguard ensures that our single-trial IAF estimates are valid and not artificially influenced by noisy data. First, we compared the number of trials in which a clearly identifiable alpha peak was detected between correct and incorrect responses. Our analyses revealed no significant differences in the average number of reliable IAF estimates across conditions (t_{115} = 1.10, p = 0.27, BF = 0.19). This result indicates that the signal-to-noise ratio—and thus the quality of the IAF extraction—is comparable between conditions, minimizing the possibility that our findings are driven by differences in data quality between correct and incorrect trials.

Then, to validate the results included in the first version of the manuscript, we replicated the tercile-based analysis and the trial-by-trial analysis by using this time the single trial estimate deriving from the analysis described at point B). Thus, we divided trials into three parts based on the prestimulus IAF, comparing performance metrics between the lowest and highest terciles. Consistent with our initial findings, we observed that both sensitivity and accuracy were significantly higher in trials within the highest IAF tercile, while no significant effect was noted on the response criterion. Subsequentially, we performed regression analyses in which fluctuations in the IAF on individual trials were used to predict perceptual performance. This analysis confirmed that a higher IAF on a given trial was associated with an increased probability of correctly detecting the target, further corroborating the robustness of our findings.

We reported these new analyses in the supplementary material:

FFT-based method for estimating single-trial IAF confirmed the pattern of results observed with the instantaneous frequency approach.

To further validate the reliability of single-trial IAF estimation, we compared the results obtained using the FFT-based method with those derived from the instantaneous frequency approach. Specifically, we applied the Corcoran et al. (2018) algorithm to the single-trial power spectra, allowing us to extract IAF values based on a robust peak detection method rather than phase-based instantaneous frequency calculations. The key advantage of this approach is its reliance on spectral features rather than time-domain fluctuations, reducing the influence of transient noise. Our analyses revealed a strong correspondence between the two methods. Specifically, trials in which the FFT-based IAF estimates were in the third terciles showed higher accuracy (t_{115} = 3.03, $p < 0.01$, BF = 7.64) and sensitivity (t_{115} = 2.77, $p < 0.01$, BF = 3.85), while maintaining the same level of criterion (t_{115} = 1.46, p = 0.15, BF = 0.29). Moreover, fluctuation in trial-by-trial FFT-bases IAF estimate impacted the probability of making a correct detection of the target (t_{115} =, $p < 0.01$, BF = 40.31). Therefore, both methods consistently indicated that higher IAF was associated with better perceptual performance, confirming the robustness of our original findings. These results suggest that different analytical frameworks can reliably capture the same underlying neural dynamics, reinforcing the conclusion that prestimulus IAF fluctuations play a crucial role in shaping perceptual outcomes.

D) Additionally, to ensure that our conclusions are not dependent solely on the single-trial approach, we replicated our analysis using a more classic method. In this approach, we computed the power spectrum by pooling all trials for each subject, separately for the correct vs. incorrect trials, and then visually inspected the spectra to identify the alpha peak. We included only those subjects for whom the alpha peak was clearly identifiable in both conditions. Subsequent t-tests confirmed that the pooled IAF was significantly higher in correct responses. We further applied the Corcoran algorithm to these pooled data, and the results were consistent with the classical method, thereby reinforcing our original observations.

We included these new analyses in the supplementary material:

IAF extracted from the pooled correct trials was higher than that extracted from the pooled incorrect trials.

Beyond single-trial analyses, we also examined whether IAF differences between correct and incorrect trials were evident when estimated from pooled data. This approach provides a complementary perspective, as it avoids trial-by-trial variability and instead focuses on broader spectral properties of the prestimulus signal. Specifically, we computed the power spectrum by pooling all trials within each response category (correct vs. incorrect) and identified the alpha peak separately for each condition. Our results revealed a systematic difference: IAF extracted from the pooled correct trials was significantly higher than that extracted from the pooled incorrect trials. This finding was consistent across subjects and was confirmed using both visual inspection of the power spectrum ($t_{100} = 3.11$, $p < 0.01$, BF = 9.80) of the spectra and the Corcoran et al. (2018) algorithm for automated peak detection applied to the aperiod-free power spectrum ($t_{81} = 2.77$, $p < 0.01$, BF = 3.73). Importantly, this pattern replicated the results obtained using the single-trial approach, reinforcing the conclusion that prestimulus IAF is a key predictor of perceptual accuracy. The consistency across different estimation methods highlights the robustness of the observed effect, ruling out the possibility that it arises from noise-related fluctuations in single-trial analyses.

E) Lastly, we conducted a trial-by-trial correlational analysis to assess the stability of the IAF estimate across the instantaneous frequency and FFT-based methods. Crucially, we found a strong correlation between the trial-by-trial IAF estimates obtained from the instantaneous frequency and the FFT. This additional analysis has been included in the supplementary materials:

The FFT-based and instantaneous frequency approaches result in comparable single-trial IAF values.

A direct comparison of single-trial IAF values obtained through the FFT-based and instantaneous frequency methods revealed a significant association between the two approaches ($r = 0.65$, $p < 0.01$, BF > 1000). This strong correspondence suggests that the two methods capture the same neural signal properties, albeit from different analytical perspectives.

Collectively, these analyses—ranging from a refined spectral approach focusing on a narrower frequency range to alternative IAF estimations methods—demonstrate that the relationship between IAF and perceptual accuracy is robust. The consistency of the IAF-related effect across multiple analytical frameworks indicates that the observed differences are not artifacts due to variations in non-alpha power or differences in signal-to-noise ratio between conditions. Instead, our findings seem to reflect a genuine neurophysiological mechanism underlying perceptual performance. We believe that these additional analyses and validations address the reviewer's concerns regarding the influence of alpha power on IAF estimation and the reliability of our single-trial measures.

That said, compared to the field addressing similar questions, the authors have already conducted extensive analyses. Given that I am joining the review process at a late stage, I wouldn't insist on these additional checks but suggest them as potential refinements or discussion points.

Thank you for your thoughtful comments and for recognizing the depth of our analyses. These additional analyses that you suggested have been particularly useful to implement, as they allow us to demonstrate that our findings are not only robust but also hold when using different methods for extracting the peak at the single-trial level. We truly appreciate your valuable input in strengthening our work.

Although addressing this issue is not strictly within my role as a substitute for Reviewer 2, I found the response to Reviewer 1's concern about the potential influence of eye movements to be overly defensive. While ICA correction removes a significant portion of instantaneous volume conduction effects, it does not necessarily eliminate the neural activity associated with eye movements. There is exciting ongoing research exploring the relationship between oculomotor behavior and alpha activity (see, for example, work by Ole Jensen's lab). If the authors wish to further investigate this, they could consider analyzing the ‘blink component’ (acknowledging that this primarily captures vertical movements) or defining a frontal electrode group for analysis using the non-eye-movement-corrected data. However, I do not insist on these additional analyses.

We thank the reviewer for this comment and for highlighting the developments in the field regarding the interplay between oculomotor behaviour and alpha activity. We agree that, while ICA correction is effective at removing ocular artifacts, it does not completely eliminate the neural signals associated with eye movements. We recognize that recent work has begun to unveil how oculomotor activity may relate to, or even modulate, alpha oscillations. In our current study, we prioritized a rigorous ICA-based correction to minimize the contribution of eye movement artifacts in our EEG analyses. Our approach was carefully designed to ensure that the observed IAF-related effects were not spuriously driven by ocular artifact. Although we do not include an extensive analysis of the blink component—primarily because such analyses extend beyond the scope and central aims of our study—we have taken the reviewer’s suggestion into account, adding a discussion note in the revised manuscript. In this note, we propose that future work could investigate whether and how oculomotor activity interacts with IAF-dependent effects or contributes independently to perceptual performance.

Moreover, follow-up studies could investigate the relationship between eye movements and IAF-related effects, as several recent and influential theories suggest a link between alpha activity and eye-related activity (Popov, T., Gips, B., Weisz, N., Jensen, O., 2023). For example, analyzing how blink-related ICA components interact with the effects we highlighted could provide valuable insights.

Reviewer #5 (Remarks to the Author):

In an EEG study of the role of alpha oscillations in visual perceptioin, the authors investigate how indivdiual and instantaneous alpha frequency (IAF), together with alpha phase, influence visual accuracy and sensitivity in a visual detection task. Their key findins is that IAF and phase jointly modulate visual perception, and the authors provide data and derive a qualitative model that lower IAF makes the impact of phase more likely. The study is well designed, uses a large sample (n =125) and sophisticated analysis methods including SDT and DDMs. Overall, the work appears as an important contribution in the ongoing debate of whether IAF modulates visual perception, including a novel view how IAF may link to alpha phase effects on visua perception.

I have been added as a reviewer after first rounds of revision, and have been asked to especially focus on a previous reviewer (#1) comment regarding the computation of instantaneous IAF, potentially being biased by 1/f noise and SNR, power confounds, and eye motion confounds. So I mainly focus on the author's reply to these concerns:

• Inst. IAF approach: The authors provided four additional analyses where they avoided computing instantanous IAF, but computed and compared IAF from power spectra of trials with correct vs. incorrect response using establised methods (Corcoran protocol, FFT methods in a fixed prestimulus interval). Specifically, they could reproduce their main findings (higher IAF has better visual accuracy/sensitivity) when estimating alpha peaks from single-trial data using the Corcoran approach. Further, the authors provide an alternative approach to remove 1/f noise using polynomial fitting. In all cases, the author could replicate their main findings. While I share the concerns of reviewer #1 that the instantaneous IAF may be partially a misnomer because frequency cannot be estimated instantenously (i.e., it needs extensive time windows for computation to achive sufficient frequency resolution to distinguish tiny frequency differences between conditions), the author provide convincing complementary analysis their their results are robust using other methods, avoiding the inst. IAF method's potential pitfalls.

• Power confounds: The authors provide additional analyses showing that power differences between correct and incorrect decision are absent or minimal (also using Bayes factors), esp.

compared to the IAF effects. (Minor: the supplemental fig. seems to lack a color bar description and a figure legend). So power confounds of the IAF effects appear unlikely.

We thank the reviewer for this valuable feedback. Moreover, we have now added a color bar description to the figure, specifying what the color represents, as this information was missing in the previous version.

• Previous reviewers also mentioned potential eye motion confounds, which were adressed previously. I do not think that eye motion confounds are specifically problematic in this study because at prestimulus time points, before any stimulus, participants have no information to alter their eye motion behavior. In principle, a link between differentiial eye motion (e.g. blinks) and IAF is conceivable, where IAF then arose from the eye motion, but this appears rather unlikely. A simple control analysis would be to compare e.g. blink frequency before stimulus onset between correct and incorrect trials.

We thank the reviewer for bringing up this point. As correctly noted, eye motion confounds were already considered in previous versions of the manuscript. We would remark here that we performed ICA-based artifact correction on all EEG datasets, carefully identifying and removing components associated with blinks and eye movements based on their characteristic scalp topographies and time courses. This procedure is a well-established approach to eliminate ocular artifacts from EEG data (Cohen, MIT Press, 2014). Moreover, as you stated, since the effects of interest concern prestimulus time points, when no visual information is yet available to drive systematic eye movements, it is highly unlikely that residual eye activity could explain the observed differences. Taken together, these steps ensure that our findings cannot be attributed to eye motion confounds.

Overall, I want to emphasize that the study is very well conducted and an important contribution.

Thank you for your thoughtful evaluation and for recognizing the rigor of our work and the effectiveness of our responses to prior concerns.

When reading the manuscripts, some additional points (none major) came to my attention, that may help to further improve the work:

Intro:

• The theoretical link between IAF and alpha phase appears plausible on first sight, because faster IAF leads to faster cycling trough optimal and non-optimal phases, so that the optimal phase is more likely hit by the stimulus. On the other hand, faster cycling also means that the oszillator stays only for a shorter period within the optimal phase, compared to a slower IAF. So the 'positive' effect of faster phase cycling could be well offset by a shorter duration within the optimal phase. This could be formalized/visualized by a phase-over-time plot.

We thank the reviewer for this insightful comment. We fully agree that faster alpha rhythms entail a trade-off between more frequent oscillatory cycles, providing increased opportunities for a stimulus to coincide with an optimal excitability phase, and shorter durations spent within each instantaneous phase. This is indeed a key conceptual point that we address more systematically in our response to your final comment, which relates to the same issue (see below). We also have now explicitly acknowledged this conceptual balance in our theoretical reasoning in the revised Discussion (see below). Briefly, we consider it unlikely that perceptual accuracy depends merely on how long the system remains within a single optimal phase. A longer single optimal phase represents an integration window of a single sampling unit, like a snapshot, where a longer duration allows for just one sample rather than multiple successive ones. Empirical evidence indicates that slower alpha rhythms favor temporal integration at the expense of segregation and temporal precision (Wutz et al., 2018, PNAS; Samaha et al., 2015, Current Biology; Cecere et al., 2015, Current Biology; Frisoni et al., 2025 Cortex). Such prolonged integration may be advantageous in contexts requiring stable perceptual grouping (Wutz et al., PNAS, 2018), but in tasks demanding rapid updating or precise temporal parsing, faster alpha cycling — promoting segregation and dynamic sampling — provides a clear advantage. Faster alpha rhythms may therefore enhance perception by enabling multiple rapid "checks" of the sensory input, allowing the brain to re-sample and refine information within the same temporal window. This interpretation anticipates the theoretical reasoning expanded in the revised Discussion, where we integrate this trade-off into a broader framework of recurrent sensory sampling and dynamic updating (see also Samaha et al., *Current Biology*, 2015).

Importantly, this reasoning also suggests that what matters may not be the absolute duration spent within an optimal phase, but rather the number of distinct opportunities the system has to align with such a phase. In other words, repeated sampling may confer a stronger benefit than prolonged single exposures, as each new cycle provides an independent chance to validate or revise the initial sensory evidence. A similar principle has been empirically demonstrated in the two-flash fusion paradigm[8], where temporal discrimination was not enhanced by longer single perceptual episodes, but instead by the opportunity to generate two separate sampling events. Consistent evidence comes from Cecere et al. [10], who demonstrated that individual alpha frequency determines the temporal window of perceptual integration, with slower alpha rhythms leading to longer integration windows and, consequently, greater temporal fusion across sensory events. Similarly, Wutz et al. [45] found that the peak frequency of alpha oscillations increased when visual task demands required temporal segregation compared with integration. Thus, a faster IAF increases not only the likelihood of "hitting" the optimal phase at least once, but also the number of distinct samples that can be accrued within a given interval. By contrast, slower rhythms, although extending the duration of each phase, risk both missing the optimal phase altogether and relying excessively on a single, potentially misleading shot of information. Taken together, this framework reinforces the view that perceptual accuracy is optimized through multiple brief samplings rather than prolonged single exposures.

We have also amended the caption of Figure 7 to clarify this reasoning, emphasizing that the duration spent within a single, even optimal, phase is not the key driver of perceptual accuracy Rather, accuracy depends on how many opportunities the oscillatory cycle has to align with an optimal phase across recurrent stimulus presentations. This clarification highlights that faster alpha frequencies enhance perceptual performance by increasing the likelihood of such optimal phase alignments within a given time window.

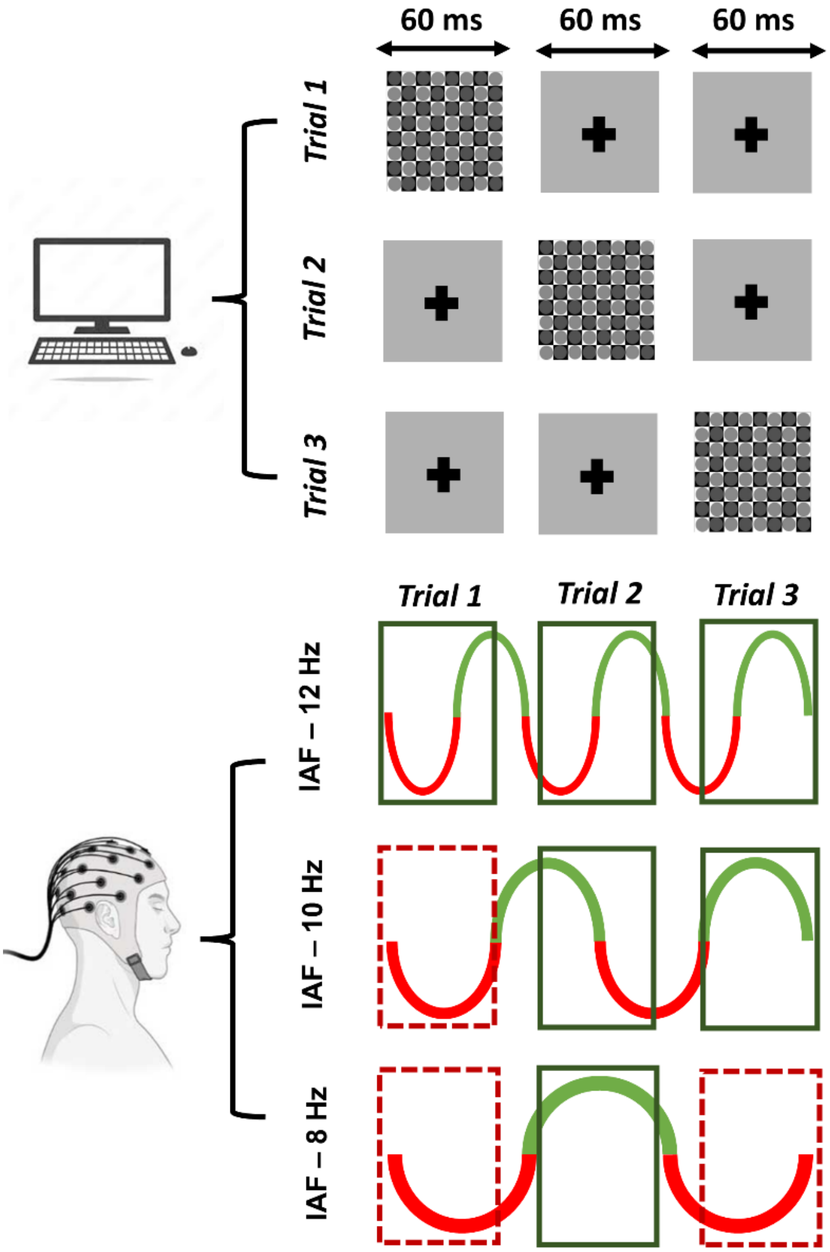


Fig 7. Alpha frequency shapes perceptual sensitivity through optimal alpha phase probability occurrence.

The mechanism that ensures a higher IAF resulting in a more accurate sensory representation relies on the increased probability of the external signal aligning with the optimal phase for stimulus perception. The rationale is that an individual with a higher alpha frequency can navigate through a wider array of phase angles within the same timeframe compared to someone with a lower alpha frequency. To illustrate the proposed mechanism, we have represented three individuals with different IAF (12 vs. 10 vs. 8 Hz). Furthermore, we have created three different scenarios in which the stimulus is presented in a different temporal moment. Green boxes represent accurate responses, while red boxes represent wrong responses. During the stimulus presentation (59 ms), individual with higher alpha individual covers more ground, encountering and embracing various phase angles, while the lower alpha counterparts traverse a narrower spectrum during the identical processing window. This difference is crucial in understanding decision outcomes. For low IAF individuals, the time point of visual stimulus occurrence hitting a specific alpha phase angle becomes a critical factor for successful stimulus processing. This is explained by the less likely transition from an unfavourable phase bin to a favourable one precisely at the time of stimulus occurrence. In essence, the limited coverage of lower alpha diminishes the probability of aligning with optimal phases during stimulus presentation, amplifying the impact of phase positioning on decision outcomes. This conceptualization highlights that perceptual accuracy would depend not on how long the system remains within an optimal phase, but on how many opportunities it has to align with one.

Results:

• Time-resolved analyses: Strikingly, sensitivity and accuracy already diverge (even though non-significantly) much earlier than the stimulus onset, already at -800ms. One would expect that at least for very early prestimulus time points (e.g. -1.5 s), IAF is not predictive of sensitivity/accuracy. Analyzing such a temporal emergence over longer time periods would serve as an important control, enhancing the specifity of a main finding.

We thank the reviewer for raising this important point. To address it, we extended our analyses to an earlier prestimulus window (–2500 to –100 ms). These additional results show that IAF predicts sensitivity and accuracy even at ~1.5 s before stimulus onset. This suggests that IAF captures an oscillatory state that can influence perceptual performance well in advance of stimulus onset. Importantly, however, when considering very remote periods (–2500 to –1500 ms), performance is indistinguishable across the considered terciles. Moreover, the predictive effect emerges more clearly as the stimulus approaches, and becomes markedly stronger and more consistent in the immediate prestimulus interval. Finally, such long-range predictive effects are not unprecedented. For example, Samaha et al. (PNAS, 2015, Fig. 4B, see below) observed that IAF predicted performance already about one second before target onset. Our findings are thus consistent with prior work showing that alpha dynamics can convey predispositional signals well before the stimulus, while gaining functional specificity in the critical prestimulus period.

From Samaha et al., PNAS, 2015

Figure Redacted

We included the presented analysis in the new version of the supplementary materials

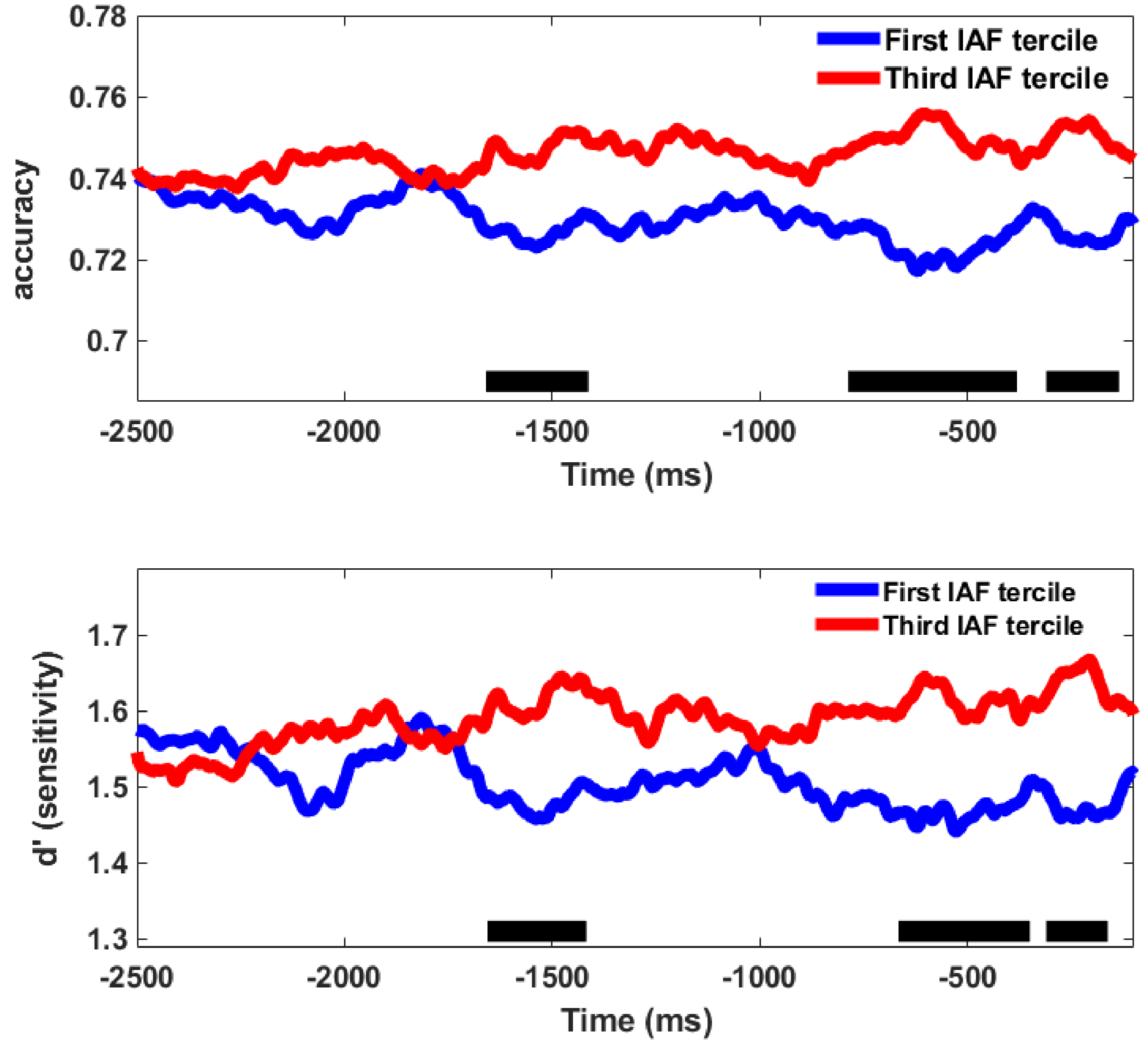


Figure S1. Time-resolved predictive effect of IAF on perceptual performance on a longer timescale

Time-resolved regression analyses showing the relationship between individual alpha frequency (IAF) and behavioral performance across extended prestimulus windows (-2500 ms; -100 ms). IAF predicts sensitivity (d') beginning ~1.5 s before stimulus onset, with effects becoming progressively stronger and more consistent closer to stimulus presentation. A similar pattern is observed for accuracy, with no differences across terciles in very remote periods (–2500 to –1500 ms), but robust predictive effects emerging in the immediate prestimulus interval.

• Alpha phase analysis: The author report alpha phase effects that occur around stimulus onset (starting at -49ms) or after stimulus onset. Thus, it appears that the phase effect is related to stimulus processing, because the prestimulus effect could well arise from temporal smearing of the time-frequency analyses into the the peri-stimulus window. So the phase effect may not arise from ongoing alpha oscillations. Further, the authors may show a (supplemental) figure of the time course of the effect.

Thank you for this useful comment. We have now included an additional figure (new Supplementary Figure S6) showing the full-time course of the alpha phase effect in the whole sample. Regarding the concern about potential temporal smearing, we would like to clarify that the reported phase effect reflects a phase-dependent modulation of perceptual accuracy - that is, behavioural performance differs as a function of the instantaneous alpha phase in the peristimulus time. We agree that temporal smearing in the time-frequency decomposition may blur the exact temporal boundaries of the estimate, but it cannot artificially produce a systematic relationship between phase and behaviour. Smearing would affect temporal precision, not the existence or direction of a phase-behaviour coupling. Importantly, when comparing the phase-related effect in participants binned by their individual alpha frequency (i.e., high and low IAF individuals), we observed significant phase effects also at time points well before stimulus onset (i.e., -685ms to -666 ms, -639 ms to -619 ms, -588 ms to -564 ms, See Fig. 4A) in the low-IAF group - that is, the group in which, according to our hypothesis, the phase-related modulation is expected to manifest more strongly. This temporal distribution further confirms that the effect does not stem from stimulus-evoked contamination, but rather reflects ongoing oscillatory dynamics influencing perceptual performance.

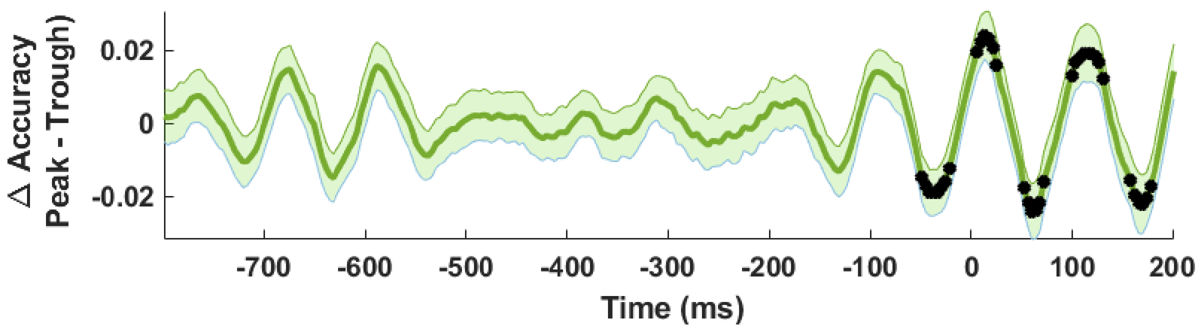


Figure S6. Alpha phase modulates perceptual accuracy.

Time course of the difference in perceptual accuracy between the two alpha phase bins (0°–180° vs. 180°–360°) across the -800 to 200 ms window. Dots indicate significant clusters where perceptual accuracy differed as a function of phase.

Significant time periods were observed both before and after stimulus onset (−49 ms to −21 ms, 5–25 ms, 52–71 ms, 99–130 ms, and 157–177 ms), indicating that ongoing alpha phase predicts perceptual performance.

• IAF x phase interaction analysis: This analysis mixes within- (phase) and between (IAF) subject levels: Showing that a phase-effect exists in low IAF individuals, but not in high IAF individuals does not allow the conclusion of a significant difference of phase effects between IAF groups (i.e. absence of evidence is not evidence of absence). This could be formally assessed with an interaction effect in this analysis. In principle, this analysis could also be computed purely within individuals, e.g. using a 2 x 2 binning approach on IAF x phase bin, showing an interaction effect. The trial-by-trial analysis of interaction effects shows that this is in principle the case, but a 2x2 approach could well visualize the effect.

We thank the reviewer for this important suggestion. To formally test for an IAF x phase interaction, we implemented an analysis specifically designed to quantify whether the effect of phase on perceptual accuracy differs as a function of instantaneous IAF. The procedure unfolded as follows.

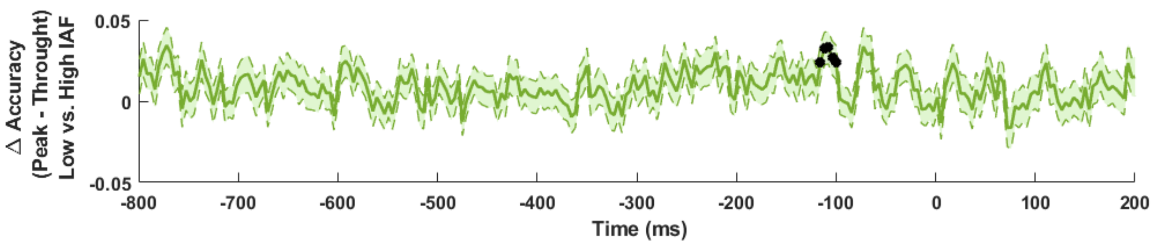
For each participant, and for each time point within the analysis window (–800 to 200 ms), we first computed trial-wise measures of phase and instantaneous IAF. Trials were then assigned to bins along two factors: (i) phase, by splitting trials into *peak* (0° to 180°) and *trough* (–180° to 0°) bins, and (ii) IAF, by dividing the trial-wise instantaneous IAF distribution into terciles, and selecting the lowest (low-IAF) and highest (high-IAF) terciles to maximize separation.

For each participant and each time point, this produced a 2 × 2 design (Phase × IAF bin). Within each cell of this design, we calculated mean perceptual accuracy. We then derived a phase effect separately for low-IAF and high-IAF trials, defined as the difference in accuracy between peak and trough bins. Finally, we quantified the interaction contrast (IC) quantified as the difference in magnitude of the phase effect between low-IAF and high-IAF trials, i.e. $|\text{Peak–Trough}|_{\text{low-IAF}} – |\text{Peak–Trough}|_{\text{high-IAF}}$. This yielded, for each participant, a time course of IC values across the entire prestimulus interval. Positive IC values indicate that the phase effect (i.e., the accuracy difference between peak and trough) was larger in low-IAF trials than in high-IAF trials, meaning that stimulus processing depended more strongly on the ongoing phase when the instantaneous alpha frequency was relatively slow. Conversely, negative IC values would indicate a stronger phase dependence when the instantaneous alpha frequency was high.

To statistically evaluate this interaction contrast, we submitted the resulting IC time courses to a cluster-based permutation test. Specifically, for each time point we performed a one-sample t-test of IC against zero, obtaining pixel-wise p-values and t-statistics. Clusters were defined as contiguous time points exceeding a pixel-wise threshold ($p < 0.05$). The cluster mass was computed as the sum of absolute t-values within each cluster. To generate a null distribution, we applied a sign-flip permutation procedure: for each of 1000 permutations, the IC values of half of the participants were randomly sign-flipped, and cluster masses were recomputed. This procedure yielded the distribution of maximum cluster masses expected under the null hypothesis of no interaction effect. Observed clusters were deemed significant if their mass exceeded the 95th percentile of this null distribution (cluster-level $\alpha = 0.05$).

This procedure thus provides a formal and rigorous statistical test for the IAF x phase interaction, directly addressing the concern raised by the reviewer. Importantly, the analysis revealed significant clusters centered around –100 ms from stimulus onset, where the interaction contrast was significantly positive. This indicates that the influence of alpha phase on perceptual accuracy was markedly enhanced when the instantaneous IAF was relatively low. This confirms that the differential pattern across groups reflects a genuine interaction, rather than the mere absence of evidence in the high-IAF group.

We have reported this new analysis in the supplementary materials as follows:



S7. IAF × Phase interaction analysis

To formally test whether the modulatory effect of alpha phase on perceptual accuracy differed as a function of instantaneous IAF, we implemented a within-subject 2 × 2 binning procedure with the factors *Phase* (Peak vs. Trough) and *IAF* (Low vs. High). For each participant and each prestimulus time point (–800 to –100 ms), trials were assigned to a phase bin based on the instantaneous alpha phase (Peak: 0°–180°; Trough: –180°–0°) and to an IAF bin based on the distribution of instantaneous frequency values (lowest vs. highest tercile). Within each bin we computed mean accuracy, yielding a 2 × 2 accuracy matrix per participant and time point. From this matrix, we calculated the phase effect separately for Low-IAF and High-IAF trials as the accuracy difference between Peak and Trough bins. To ensure that the analysis captured the strength rather than the direction of this effect, we took the absolute value of these differences. The interaction contrast (IC) was then defined as the difference in the magnitude of the phase effect between Low-IAF and High-IAF trials:

$$IC = |\mathrm{Accuracy}_{\mathrm{Peak}} - \mathrm{Accuracy}_{\mathrm{Trough}}|_{\mathrm{Low\ IAF}} \; - \; |\mathrm{Accuracy}_{\mathrm{Peak}} - \mathrm{Accuracy}_{\mathrm{Trough}}|_{\mathrm{High\ IAF}}$$

This procedure produced, for each participant, a time course of IC values across the prestimulus interval. To statistically evaluate the IC, we performed a cluster-based permutation test. At each time point, one-sample t-tests against zero were computed across participants. Suprathreshold points ($p < 0.05$) were grouped into clusters, whose mass was defined as the sum of absolute t-values. The significance of observed clusters was assessed against a null distribution obtained by a permutation analysis (1000 iterations), in which the sign of IC values was randomly inverted for half of the participants. Clusters exceeding the 95th percentile of this null distribution were considered significant (cluster-level $\alpha = 0.05$). Crucially, the modulatory effect of alpha phase on accuracy was significantly stronger for trials with low instantaneous IAF compared to those with high IAF during the prestimulus period at around -100ms.

• ITPC analysis: Again, the question is whether correct-incorrect contrast differs between high and low IAFs. This interaction effect is not tested directly.

We thank the reviewer for this thoughtful comment. We fully agree that a crucial question concerns whether the correct–incorrect ITPC contrast varies as a function of individual alpha frequency (IAF). In principle, this could be examined through a fully within-subject 2 × 2 design (IAF × accuracy) as we did in the previous analysis. However, this approach is not statistically appropriate for ITPC measures in our specific context. ITPC estimates are computed as the magnitude of the complex average of single-trial phase angles and are therefore highly sensitive to trial count and balance across conditions. Reducing the number of trials per condition, as would be necessary when subdividing each participant's dataset simultaneously by IAF terciles and response accuracy (correct vs. incorrect), introduces substantial estimation noise. This degradation of reliability is not a minor practical issue: the ITPC metric becomes biased and unstable when based on small or unequal sample sizes, since it measures the consistency of phase alignment across trials rather than a central tendency that can be corrected by averaging (Cohen, MIT Press, 2014). Thus, a purely within-subject 2×2 design would have risked producing spurious or uninterpretable results driven by differences in trial number rather than genuine oscillatory effects. Moreover, unlike the previous analysis, where trials were approximately balanced across instantaneous phase bins (Peak vs. Trough), the present contrast inherently suffers from severe cell-size imbalance, because incorrect trials are far fewer than correct trials (average proportion 3:7). This asymmetry would further bias ITPC estimates and inflate variance in a within-subject 2×2 (IAF × Accuracy) design, making any apparent interaction difficult to interpret as a genuine oscillatory effect rather than a trial-count artifact.

To overcome this limitation while still addressing the reviewer's concern, we opted for a between-group comparison of the correct–incorrect ITPC contrast between low- and high-IAF individuals. This approach ensures that each ITPC estimate is based on a sufficiently large and balanced trial set, preserving the reliability and interpretability of the phase-locking measure. Importantly, this analysis directly tests whether the strength of phase alignment associated with accurate perception differs as a function of IAF, thus addressing the interaction question in a statistically sound way. Therefore, we tested the directional hypothesis that the correct–incorrect ITPC contrast would be larger in low-IAF than in high-IAF individuals.

The results confirmed this prediction and provided formal evidence for an interaction between IAF and ITPC: the phase consistency difference between correct and incorrect trials was significantly greater in low-IAF compared to the high-IAF participants in the high-alpha low-beta band in the peristimulus window. In sum, this analysis represents the most statistically robust way to evaluate the interaction between IAF and ITPC, given the constraints inherent to the metric itself.

We included this additional analysis in the supplementary materials:

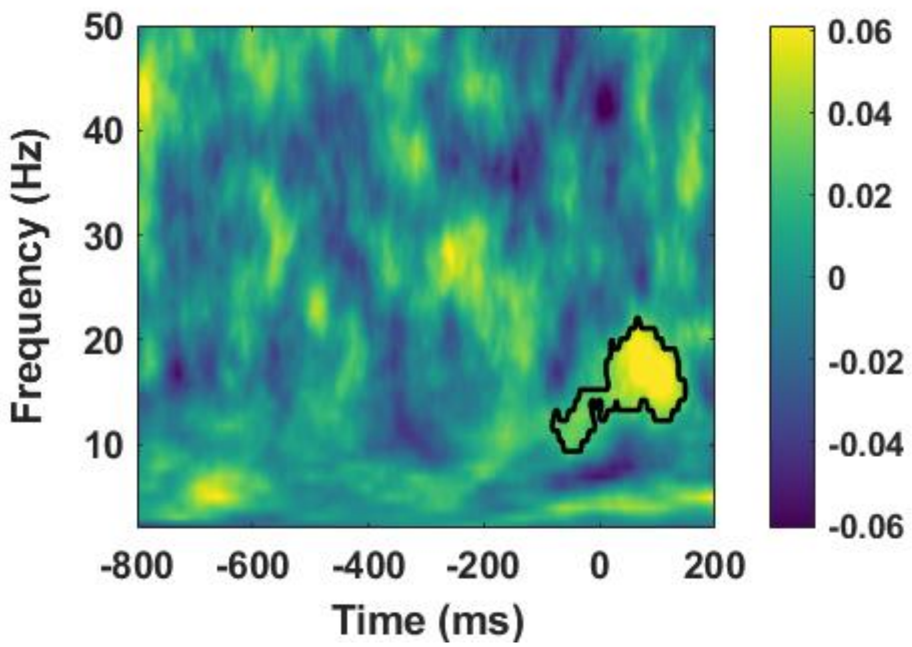


Figure S8. IAF modulates the correct–incorrect ITPC contrast in the peri-stimulus window. Time–frequency map of the between-group difference in the phase-locking contrast between correct and incorrect trials (ΔITPC), calculated as Low-IAF$_{\Delta ITPC}$ minus High-IAF$_{\Delta ITPC}$. Difference in trial counts for correct vs. incorrect responses were controlled via stratified resampling (see Methods). Black contours mark clusters significant under a cluster-based permutation test (1000 permutations). Warm colors indicate a larger ΔITPC in Low-IAF than in High-IAF group. A significant positive cluster emerged in the high-alpha/low-beta band (~9–16 Hz) during the peri-stimulus period (−80 to +150 ms), indicating that the phase-locking difference between correct and incorrect perception was stronger in Low-IAF than in High-IAF participants.

Discussion:

• Fig. 7: Even though the model is convincing at first sight, it makes a critical assumption: For accurate visual perception, the optimal phase must be hit, independent for how long stimulus processing falls into this phase. Under this assumption, higher IAF will indeed increase the likelihood that a stimulus at least briefly touches the optimal phase due to rapid cycling. On the other hand, the amount of time a stimulus is processed during the optimal phase could also be critical for accurate perception. Here, a lower IAF may increase the probability for a high optimal-phase duration. Thus, these two effects might cancel each other. Only a quantitative model could distinguish these cases.

We appreciate the reviewers' insightful comments pointing to this fundamental issue: the trade-off between faster IAF providing more opportunities to "hit" the optimal phase versus slower IAF allowing longer durations within that optimal phase. We agree that this is a crucial aspect of the theoretical framework, and we believe it opens an important line of reflection that goes beyond the purely formal description of oscillatory cycles.

From our perspective, it is unlikely that perceptual accuracy is maximized simply by prolonged staying in the optimal phase window during a single cycle. Sensory systems are designed to operate in a highly dynamic and stochastic environment, where rapid sampling and the possibility to "double-check" incoming evidence play a fundamental role. In this sense, the advantage of a higher IAF may not lie in the absolute amount of time spent in the optimal phase during one oscillatory cycle, but rather in the *increased number of opportunities* to realign perception with optimal phases across successive cycles. Each cycle can be conceived as a "sampling unit," and having more cycles within a given temporal interval increases the probability that a stimulus will be processed not just once, but repeatedly, at or near an optimal phase.

This logic is consistent with recent views that perception is not a single-pass phenomenon but is continuously updated through recurrent sampling of sensory input. From this perspective, the critical factor is not how much information is extracted in a single "shot," but rather how many times the system can re-sample and possibly revise or corroborate the information extracted. The first extraction may already carry the core of the information, and subsequent accumulation within the same phase interval may be subject to rapid saturation. That is, once the system has acquired the essential features of the input, lingering in the same phase may not yield proportionally more accurate or richer representations. In contrast, engaging in multiple shorter samples across different cycles allows the system to test, refine, and potentially falsify the initial evidence, leading to a more robust perceptual inference.

For example, empirical support for this view comes from studies using temporal resolution paradigms, such as the 2-flash fusion task (Samaha et al., *Current Biology*, 2015). In that work, perceptual accuracy was shown to depend critically on whether the visual system could generate two distinct sampling events, rather than on the mere duration of a single perceptual epoch. In other words, having two separate samples of the stimulus, even if each is brief, provides more reliable information than a longer single exposure. Consistent evidence comes from Cecere et al. (*Current Biology*, 2015), who demonstrated that individual alpha frequency determines the temporal window of perceptual precision, with slower alpha rhythms leading to higher proneness to the illusion. This finding fits with our interpretation that perceptual accuracy benefits from the multiplicity of phase-aligned sampling opportunities, rather than from extended duration within a single optimal phase.

Crucially, a higher IAF not only enables more frequent sampling opportunities, but also statistically increases the likelihood that a stimulus will fall into the optimal phase at least once. With slower oscillations, there is a greater risk that the temporal alignment between stimulus onset and phase position will "miss" the optimal window altogether, such that the stimulus is processed under suboptimal conditions. In this respect, faster rhythms provide a double advantage: (i) they multiply the opportunities to engage in iterative sampling, and (ii) they reduce the risk of complete misalignment between stimulus timing and phase position.

This argument also helps reconcile why faster IAF could be beneficial despite shorter optimal-phase intervals: the multiplicity of phase-aligned samples offsets the reduced duration of each, and may even provide an adaptive advantage. Indeed, from the perspective of predictive coding frameworks, perception benefits not from passively accumulating redundant information, but from actively updating the internal model whenever the sensory stream provides an opportunity to do so. Multiple phase-locked samples provide precisely this opportunity for dynamic updating.

We agree that future work could empirically test this hypothesis using computational modeling approaches. For example, models could simulate perceptual accuracy under conditions where information is either accumulated continuously during one prolonged optimal-phase epoch versus conditions where multiple shorter, temporally separated optimal-phase epochs provide independent sampling opportunities. Such simulations would allow quantifying the relative gain of duration versus multiplicity, and to determine whether a saturation effect indeed limits the benefit of longer single-phase exposures.

While future modelling work could help to formally test this hypothesis, we want to further emphasize that existing empirical evidence already provides strong support for our hypothesis. Studies consistently show that slower alpha rhythms are associated with longer temporal integration windows and reduced temporal precision, a pattern that directly contradicts the notion of enhanced evidence accumulation within a prolonged optimal phase. Apart from the Samaha et al. (*Current Biology*, 2015) and Cecere et al. (*Current Biology*, 2015) studies (for a review and metanalysis on the consistency of these effects see Samaha and Romei, 2024 J Cog Neurosci), similar conclusions have been reached by Wutz et al., (*PNAS*, 2018) and by Ronconi et al. (*PNAS*, 2017), who showed that slower alpha oscillations expand the perceptual integration window but at the cost of worsen temporal resolution. Together, these converging findings indicate that slower alpha rhythms reflect a less dynamic and less adaptive sampling regime, making the perceptual system more prone to interpretative errors rather than supporting a more efficient accumulation of sensory evidence.

In conclusion, while we acknowledge the conceptual trade-off emphasized by the reviewers, our data leans towards the interpretation that multiplicity of sampling and increased probability of phase alignment are more advantageous than prolonged single exposures. This view also aligns with empirical evidence on the benefit of recurrent sensory sampling, such as the 2-flash fusion findings by Samaha and colleagues, but also situates the IAF–phase relationship within a broader computational principle: perception as an iterative process of hypothesis testing rather than one-shot accumulation. We are grateful for this opportunity to clarify our point of view that we integrate in the discussion:

Importantly, this reasoning also suggests that what matters may not be the absolute duration spent within an optimal phase, but rather the number of distinct opportunities the system has to align with such a phase. In other words, repeated sampling may confer a stronger benefit than prolonged single exposures, as each new cycle provides an independent chance to validate or revise the initial sensory evidence. A similar principle has been empirically demonstrated in the two-flash fusion paradigm[8], where temporal discrimination was not enhanced by longer single perceptual episodes, but instead by the opportunity to generate two separate sampling events. Consistent evidence comes from Cecere et al. [10], who demonstrated that individual alpha frequency determines the temporal window of perceptual integration, with slower alpha rhythms leading to longer integration windows and, consequently, greater temporal fusion across sensory events. Similarly, Wutz et al. [45] found that the peak frequency of alpha oscillations increased when visual task demands required temporal segregation compared with integration. Thus, a faster IAF increases not only the likelihood of "hitting" the optimal phase at least once, but also the number of distinct samples that can be accrued within a given interval. By contrast, slower rhythms, although extending the duration of each phase, risk both missing the optimal phase altogether and relying excessively on a single, potentially misleading shot of information. Taken together, this framework reinforces the view that perceptual accuracy is optimized through multiple brief samplings rather than prolonged single exposures.

Minor

• Line 335: The polar plot should only be presented describtively, but phases should not be tested again because they arise from the already significant cluster (double dipping).

We thank the reviewer for pointing this out. We fully agree and would like to clarify that the polar plot was included for descriptive purposes only. The p-value reported there refers to the original cluster-based permutation test, and no additional statistical test was performed on the phase values (i.e., no double dipping).

• Line 434: Reference to Fig. 7 may make more sense?

We thank the reviewer for the suggestion. We agree that the reference to Fig. 7 is more appropriate, as the mechanism is better illustrated there. We have revised the manuscript accordingly.

The authors responded extensively and convincingly to my previous comments, adding some important specifications to their model and adding important control analyses:

We thank the reviewer for the careful re-evaluation of the manuscript and for the positive assessment of the additional analyses and clarifications. We respond point-by-point below.

• As stated in my previous comments, I find the control analyses in response to reviewer #1 comments convincing.

We thank the reviewer for acknowledging that the additional control analyses are convincing.

• The authors refined their model to link IAF and phase, accounting now for the possibility that duration within an optimal phase may play a role. I agree with their view that rapid perceptual sampling is more important for the system in dynamic environment than the sampling duration of a stimulus in the optimal phase – sensory systems need to represent change, not constancy. On the other hand, if only the speed of cycling is crucial, why does the brain use a comparably slow oscillation around 10 Hz for perceptual sampling, and not, say, 100 Hz? There must be some upper limit in the perceptual sampling frequency, which arise from internal (e.g. “over-sampling” leads to unnecessary information load) or external factors (e.g. the temporal statistics of usual environment do not require fast sampling).

We appreciate the reviewer’s thoughtful theoretical comment. We agree that if perceptual performance depended exclusively on sampling speed, one might ask why perceptual sampling operates in the alpha range rather than at much higher frequencies (e.g., gamma). As the reviewer suggests, the emergence of alpha-range sampling is likely shaped by a combination of intrinsic and extrinsic constraints. In line with the reviewer’s point on “oversampling,” increasing sampling rate may inflate redundant updates without proportional gains in information, while increasing metabolic cost and potentially destabilizing recurrent inference. Moreover, the dominant timescales of natural visual dynamics and active sensing (e.g., fixation-to-fixation updates) may not require ultrafast sampling, making an alpha-range rhythm a plausible compromise between temporal resolution and reliability. Finally, it is notable that alpha activity can emerge “naturally” in computational models of visual cortex: for instance, Alamia and VanRullen (2019) show that communication delays in a predictive-coding model give rise to alpha-like reverberations and traveling waves during visual stimulation. We agree that these ideas remain hypotheses and should be tested directly in future work. To reflect this point, we added the following brief clarification to the Discussion:

“Taken together, this framework aligns with the view that perceptual accuracy is optimized through multiple brief samplings rather than prolonged single exposures. We hypothesize that the alpha range may reflect a timescale compatible with circuit integration and long-range communication delays, which could increasingly hamper coherent sampling at much higher frequencies. In line with this, computational models of the visual system incorporating realistic conduction delays naturally generate alpha-like reverberations and traveling waves (Alamia & VanRullen, 2019)."

- Further, the authors performed time-resolved analyses of the predictive IAF effect on perceptual performance that extended the time window to much earlier time points. This analysis shows that the effect emerges much early around -1.5s prestimulus, but not at even earlier time points. This is an important control analysis and finding.
- Finally, the authors convincingly argue and provide a control analyses showing that a temporal smearing explanation of the phase effect shortly before stim onset is unlikely.
- Finally the author performed a formal test of the interaction effect of IAF x phase as suggested. This analysis confirmed an interaction shortly before stimulus onset. I agree that a similar within-subject analysis of ITPC for IAF x accuracy interaction is impossible if cells of the 2x2 design have unequal trial numbers, but the authors show with a mixed analyses that such a prestimulus interaction is indeed present in their data.

We thank the reviewer for suggesting these helpful analyses and for finding our interpretation convincing.