

Distinct distributed neural dynamics predict pallium-dependent social approach

Corresponding Author: Dr Lilach Avitan

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

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

In this manuscript, the authors investigate the neural basis of social approach behavior at cellular resolution in zebrafish. They report a brain-wide neuronal activation pattern that precedes the movement of the focal fish, characterized by increased activity in a small subset of telencephalic neurons, along with reduced activity in the midbrain and hindbrain. This pattern appears to be specific to social attraction or engagement. Overall, I found the study highly compelling and the results both interesting and relevant. The authors introduce a novel and thoughtfully social assay that effectively captures social interactions with a real conspecific. This approach holds great potential for future studies exploring brain-wide dynamics in diverse social contexts.

The manuscript is clearly written, the figures are well-prepared and effectively support the findings, both in the main text and supplementary material, and the experiments are carefully controlled.

That said, I do have a few minor comments, which I outline below:

Throughout the manuscript, the telencephalon is referred to as the forebrain. While the telencephalon is indeed a part of the forebrain, along with the diencephalon, the analyses of these two brain regions were performed separately, as shown in Figure 3A. Therefore, it is inaccurate to refer to the telencephalon as the forebrain. This should be corrected throughout the text: the term “forebrain” should be replaced with “telencephalon,” as it is the telencephalon the authors are actually referring to.

Additionally, in lines 329–330, the text states: “These forebrain neurons were primarily located in the pallium (dorsal forebrain, Figure S5I).” However, the pallium and subpallium are subdivisions of the telencephalon, not of the forebrain. This should also be corrected for anatomical accuracy.

Lines 310–312: In this section of the discussion, the authors state that the full complexity of social cues was preserved. However, based on the description provided in the Methods section, it appears that only visual cues were accessible during the social assay. The fish were placed in separate compartments, preventing the exchange of chemical or mechanosensory cues in the water. This point should be clarified and the text revised accordingly to accurately reflect the sensory modalities available during the assay.

Perhaps I missed this information in the Methods section, but what was the imaging depth during data acquisition? Was the entire brain imaged from dorsal to ventral? I ask this in relation to the acquisition of the neuronal activity in the diencephalon. In Figure 3A, the diencephalon is depicted as a very small grey area that does not represent the size or entire location of this brain region. Anatomically, the diencephalon lies beneath the optic tectum and is a considerably large structure. Was the diencephalon fully captured? This point is unclear in the manuscript, and the representation in Figure 3A is somewhat misleading. A lateral view of the brain would provide a more informative depiction of the imaged areas.

If the diencephalon was fully acquired and does not show any distinct activation pattern, it would be very interesting to include this topic in the discussion and propose potential hypotheses as to why this brain region, typically involved in the expression of social behavior and considered part of the social behavior network, does not respond in a social affiliation assay.

Still in the discussion, it would be interesting to go beyond description and explore why more dorsal areas of the telencephalon are active. Which specific brain nuclei are located in these regions? Could this activity be related to the reward system?

Minor comments:

Line 395: The text states, "All details of the control experiment (Figure S6G,H)," but there is no Figure S6. The last figure in the supplementary information is Figure S5. This should be corrected.

Lines 338 and 346: Both paragraphs begin with similar phrases, "Our results uncovered" and "Our data revealed", which is somewhat repetitive. Consider rephrasing one of them.

Figure panels: In some figures, such as Figure 1I and others, the panels are divided into top, middle, and bottom sections. For clarity, it would be helpful to label these subpanels directly in the figure as I, I', and I'', for example, or identify them in some other way.

Caption of Figure G: The caption states, "...showing that approach vs. non-approach distinctions emerged earlier for long sequences." However, the comparison here is not approach vs. non-approach, but rather is approach and not approach. This should be revised.

Reviewer #2

(Remarks to the Author)

In this paper, Lifshitz et al., describe a novel assay for monitoring at the same time behavioral responses and neuronal activity, via calcium imaging, of a zebrafish larva in the presence of a conspecific. They used this setup to identify patterns of neuronal activity correlating with approaching of the imaged fish toward a freely swimming conspecific larva. The assay is novel and potentially very useful for studying the neuronal mechanisms underlying social behavior in zebrafish. Moreover, I think that the experiments were well designed and mostly support the authors' conclusions. However, I find the work extremely correlational and lacking any supporting evidence that the observed patterns of neuronal activity, or any of the neurons active during approach behavior, are actually required for social interactions. It is my opinion that in the current state the manuscript does not provide new insights into the neuronal mechanisms of social behavior, and should not be accepted for publication before a major revision. Please, find below my specific comments.

Major

1 - Why did the authors use prey capture behavior in freely moving fish to compare turning angles in head-fixed larvae to the head-fixed preparation? Wouldn't it be more appropriate to use turning toward conspecific in freely moving fish for the comparison and validation of the tail movement quantification method? Are not turning parameters expected to be different in the two contexts (i.e., J-turns during hunting?)

2 - The authors stated that there is an effect of position of conspecific fish on turning angle of the focal fish (figure 1G), and that the effect is not present in the dark (Figure 1H). The effect is quite small and the authors did not perform statistical test to support these two conclusions. This should be done.

3 - The authors claim that distinct neural activity preceding approach vs. non-approach movements emerged only when focal fish observed a real conspecific (lines 298-299). The difference is small, suggesting that the neuronal activity pattern observed may mostly underlie activity related to a generalized approach response, rather than one specific to social interactions. The fact that activity of only a small group of neurons is responsible for the difference between generalized and social-specific approach could be used to identify and better characterize these neurons. However, this detailed analysis was not performed and, as a consequence, the manuscript lacks any data related to identity of the neurons responsible for social interactions. More effort should be put to better characterize the difference between neuronal activity correlating specifically with social interactions.

Moreover, I'm not convinced that there is an actual statistically significant difference in the midbrain activity in Figure S5G. Why did the authors perform one-sided t-test instead of a two-sided one?

Finally, I think that data supporting the presence of neuronal activity correlated specifically with social approach is a very important piece of evidence for this manuscript, which should not be relegated to a supplementary figure.

4 - The major shortcoming of this work is that it is only a correlational study. There is no evidence that the neuronal activity pattern observed or any of the neurons active during an approach are required for social behavior. The authors should perform studies to test the effect of ablation or optogenetic perturbation of the identified neurons on approach toward conspecific fish. The following are very good examples of this type of approach: Vladimirov et al. Nat. Methods (2018) <https://doi.org/10.1038/s41592-018-0221-x>; Kappel et al. Nature (2022) doi: 10.1038/s41586-022-04925-5

Minor

1 - It's not clear what brain regions the authors consider to be part of the forebrain in their Ca²⁺ imaging analysis, since they separated forebrain and diencephalon, which is part of the forebrain. Did they mean telencephalon instead of forebrain? If so, the term should be corrected.

2 - Figure 2E. Why are most of the points representing conspecific angular position clustered between -100 and 0 degrees? The distribution in Figure 1E doesn't look that skewed.

Reviewer #3

(Remarks to the Author)

The authors have implemented a social interaction setup where a "focal" fish is head restrained and tail free, and a freely moving conspecific is placed at a distance (separated by a barrier). Using this setup, the authors have reproduced previously published observations that fish will preferentially reorient in the direction of conspecific motion. By recording from a fraction (10,000 neurons) of the "focal" fish brain at a time, the authors find neural activity patterns that are correlated with movement toward conspecifics.

While I found the subject of the manuscript to be interesting, there are significant issues with data interpretation and the overall validity of the conclusions. The advantage of this experimental setup over previous publications is unclear. The Methods section lacks essential details.

On methodology:

1. The advantages, as well as the implementation of the experimental setup, are both unclear.

1a. The experimental setup is not sufficiently described, particularly regarding the physical properties of the barrier separating the fish. Critical details such as the material's structure, thickness, refractive index, and how it might affect visual perception at various angles are missing.

1b. What is the advantage of this system? This is arguably a more simplistic and less sophisticated setup than a VR system. VR for head-restrained larval zebrafish has been implemented in multiple labs, including closed loop VR. Given that the authors have deliberately removed olfactory or somatosensory information, what can this offer that a visual display cannot?

2. The description of how neural activity was recorded is vague and lacks important technical details.

2a. "Subsequent to the functional imaging, the entire brain structure (300-350 imaging planes, 1 μ m apart) of several focal fish (n=6) was scanned for structure." – What does "scanned for structure" mean? Is 300-350 imaging plane collected per fish, or summed across all fish? At what temporal resolution? For how long? For how many sessions per fish?

2b. Previous to this statement, the authors write "Simultaneously, volumetric two-photon imaging (Nikon objective, CFI $\times$ 16, 0.8 NA) captured neural activity over a 705 $\times$ 352 $\times$ 70 μ m area at 3 volumes/sec." – This is clearly not the entire brain volume. So then what did the authors actually choose to image for a given animal? How many biological replicates exist for each brain region?

2c. In the Results section, the author separately state that they "record activity from over 10,000 neurons across the brain of the focal fish" – this is only ~10% of the larval brain, which region was recorded from each fish? How long was the recording? How many sessions were recorded per fish?

3. 16 dpf brains are registered to a 6 dpf reference brain? Given the size and structural difference between the brains at these developmental ages, this can only give crude anatomical annotations at best.

On data analysis and interpretation:

1. The authors conclude based on the topographical organization of activity in the tectum that "This confirms that the focal fish observed the natural social cue and encoded its physical properties, such as position and direction of motion, in tectal neural activity." This is a somewhat trivial statement. The retinotopic organization of activity in the optic tectum has long been established. Any object (moving or not moving, social or non-social) will be encoded topographically in the tectum.

More importantly, the authors have not demonstrated that any of the behavioral or neural results are specifically social (see #3 below)

2. The statement in the abstract "We demonstrate that zebrafish approach behavior is jointly shaped by spatial and temporal information from conspecifics rather than by these features acting independently" is also expected given existing publications. It's known that zebrafish respond to the joint position and motion information of conspecifics. Zebrafish also responds jointly to the position and motion of many other stimuli (e.g. prey, predators, objects, grating, etc.).

3. The authors claim that the "time lag between fish movements was significantly shorter than expected by chance", but the right control is not chance but stimulus motion that do not resemble fish movement, which has been demonstrated in previous studies (e.g. Kappel et al 2022), but not in the present study.

4. One of the major issues in this manuscript is the vague definition of "approach movements," which are classified as any movement toward the conspecific. This definition creates a potential bias, as movements toward a conspecific may be statistically different from those directed elsewhere. Without careful characterization of the kinematic parameters (e.g., bout speed, inter-bout interval, duration, type), it is difficult to disentangle whether observed neural differences are a result of movement patterns, sensory representation, or social approach. The lack of detailed analysis in this regard significantly weakens the authors' conclusions about the neural basis of social approach.

For example, the authors only present a single bar plot suggesting the "predicted tail angle" is not different between approach or non-approach movements. Imagine a very simple scenario, where the IBI of consecutive "approach movements" is significantly lower than the IBI of consecutive "non-approach movements". Differences in movement rate correlate with differences in global brain activity, something that has been demonstrated across species. Without taking this into account, the authors would then misinterpret any observed difference as reflective of social processing.

5. A number of conclusions are drawn from the so-call “conspecific sequences” (e.g. the “focal” fish is claimed to respond preferentially to the early phase of these sequences). What exactly are these sequences? Given the geometric constraints of the environment and movement patterns of zebrafish, one could imagine that these are mostly just the fish turning around at each end of the chamber. During the “early phase” of these leftward or rightward sequences, the conspecific is simply positioned closer to the optimal response region of the “focal” fish.

Furthermore, the conspecific seems to have a spatial preference for a particular zone of the outer chamber, which further complicates the analysis. It’s unclear what is the cause for this, and whether it is related to the presence of the “focal” fish at all.

6. The authors go on to state that: “our results indicate that social movements in zebrafish arise from long-timescale neural computations, linked to the time elapsed between a change in conspecific direction of motion and focal fish movement onset.” There is no clear support for this conclusion.

Multiple previous publications, e.g. Stednitz et al 2018 or Dunlap et al 2021, have already demonstrated that the forebrain responds to conspecifics. Thus, given that approach movements are correlated with angular position of the conspecific, the simplest explanation for much of data is that forebrain activity is modulated by the position of conspecific.

Imagine that a forebrain neuron is simply tuned to the angular position of the conspecific, peaking around 50 degrees, where the fish is most likely to turn towards the conspecific. Given the movement statistics of larval zebrafish, a conspecific is likely to spend several seconds (at least) within this angular range on each traversal through the environment. In this case, even if induction of turning towards to the conspecific is purely probabilistic (i.e. having nothing to do with evidence accumulation), when averaged over many movements, one would expect elevated forebrain activity for several seconds preceding an “approach movement”.

7. Other possible explanations for persistent activity differences before a given movement, such as the statistics of the preceding set of movements or arousal, are not considered at all.

8. Throughout, the authors have attempted to correlate neural activity to behavior. To claim that the observed activity patterns “mediate” or drive social approach would require functional manipulations, which have not been performed.

Altogether, a number of basic analyses and controls are missing. The presentation is confusing, several key interpretations are questionable, and experimental details are missing. This is not to say that valuable insight cannot be gained from the data that the authors have produced with more in depth and carefully considered analyses in the future.

Reviewer #4

(Remarks to the Author)

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

Reviewer #5

(Remarks to the Author)

In this manuscript, Lifshitz et al describe a behavioral assay for late-stage zebrafish larvae, where a head-fixed fish is able to respond to a conspecific swimming in an enclosure within their field of view. This assay allows for brain-wide functional imaging, and the authors describe a multi-regional activity pattern that is predictive of the movement direction in the fish observing conspecific swimming. This is a novel behavioral assay for head-fixed larval zebrafish, and they quantify some interesting neural activity patterns of the observer fish in response to the position/motion of the conspecific swimming in an enclosure around the animal. This paper is tackling an interesting topic, takes a careful approach to distinguish it from previous work on this topic, and is written in a clear fashion. However, I have some concerns about the methods, data presentation, interpretation of results, and integration with the published literature. I detail these points below, and encourage the authors to address these points in a substantially revised manuscript.

Major points

1) The authors state that 2-week old fish approach conspecifics, but this is debatable. While some groups claim there is social this approach by 2 weeks (refs 16,19), others find this attraction only emerges at 3 weeks (for instance, Larsch & Baier 2018). This makes the choice of 2 week fish problematic, but I understand why it is technically more challenging to image older zebrafish. That being said, the authors appropriately discuss how this behavior is variable across fish and not very robust at this age (30% do not show it here).

To validate this assay better in these animals, especially in light of the variability amongst individuals, I would like to see a comparison of the head-fixed and freely-moving social attraction in individual animals. How does the distance between fish in the freely-moving vs head-fixed cases (and subtended visual angle of those conspecifics) compare when looking at approach movements? This assay would be better justified if the authors could show that the most social of 14dpf animals in the free-moving condition are also the most social in the head-fixed condition. Furthermore, in the head-fixed prep, does the direction of movement for the focal fish and the conspecific match? Or, is it just that the focal fish moves in the direction of the side where the conspecific is, regardless of their movement direction?

2) The tail tracking algorithm is quite nice, and has many applications – probably best for hunting assays (Fig S2). But, the behavioral tracking model appears to be trained from freely-moving fish (filming fish from the top), yet applied to head-fixed

fish (filming fish from the bottom of the dish). Therefore, the authors should apply a sign-inversion to tail angles recorded from the bottom, where a left-turn from below would be a right-turn from above. Without this correction, the slight attraction reported here (Fig 1G) could in fact be a slight avoidance. This avoidance would be consistent with previous work in 2-week zebrafish (Larsch & Baier 2018; Harpaz et al., 2021), but would have a very different interpretation from the framing of this manuscript. Please clarify whether this was done in the methods.

3) The head-fixed assay is clearly useful for the neural access it provides, but it causes fish to behave in an unusual manner that was not discussed enough. For instance, are there statistical considerations for the > 10-fold difference in movement rates between the head-fixed observer (<0.1 Hz) and freely-moving fish (>1 Hz)?

Furthermore, is the freely-swimming fish showing any restraint stress? How long do the animals habituate before the behavior/imaging?

Finally, is the Poisson model applied here (Fig S3) using these disparate movement rates? This would be important to model appropriately.

4) A key conclusion of this paper is about the long neural timescales seen. Cosine distances and neural activities are clearly different for the entire 10s pre-stimulus period in most plots. This is quite interesting, but difficult to interpret without knowing when these changes began. Can the authors plot a longer pre-stimulus period to find out when neural activity diverges between the type of upcoming movements? This is a key question that is central to the authors claims, but not examined sufficiently. If there are "long timescales" of activity predicting social engagement, what is that timescale? It is clearly longer than 10s.

Also, this type of >10s mid-/hind-brain activity predictive of upcoming movement direction in larval zebrafish has also been observed in a non-social behavior (Lin...Vaziri, Cell 2020), which was not cited.

5) For the observed decreased in midbrain neurons: is this observed across the whole midbrain, or is there selective decrease/increase in neurons selective to the social object (ie: those cells identified in Figure 1I)?

What are the anatomical connections here that link increased forebrain activity to suppressed midbrain/hindbrain? The authors claim there are directed projections from the forebrain to the midbrain and cite ref. #38 (Kunst et al), but there is no evidence for this. That paper states "For example, we found that the forebrain together with the preoptic region and the posterior tuberculum comprises a distinct module anatomically separated from the mid- and hindbrain, which seem to be much more strongly interconnected". What are the authors referring to? Do they have evidence that forebrain excitation precedes midbrain inhibition on a single-trial basis?

Minor points

1) The conspecific angular position on the observer's tectal map is a nice validation of social vision (Fig 1I), but the authors also say they see direction selectivity (line 117). Please show this map too - does this have any spatial order?

2) For the argument about movement sequences of duration 9.5s : Given that the average movement rate is ~5 movements per minute (Fig S1B), wouldn't that typically just be one movement? Most of the examples in Fig 2H are just showing 1 movement per "sequence". How is this considered a sequence?

3) the Y-axes in many plots are not matched to allow for fair comparisons (for instance, Fig S4C, Fig S4G, Fig 3B, Fig 3C,D,F,

4) when neurons are averaged or used for population-level analyses (decoding, cosine dist.), the authors chose a variety of different neurons #s (100 in some cases, 300 in others, 20 in other conditions). Why was this done? Please justify these choices in the methods section, or be consistent throughout.

5) I was very surprised to see no description or discussion of neural activity in the dorsal thalamus – the main brain region identified for social attraction in zebrafish (see Kappel et al., Nature 2022, citation #28 here). As this is the other major study of neural imaging in response to social motion in zebrafish larvae/juveniles, I found this to be a notable omission. If the authors did not see biological motion-sensitive neurons in the midline dorsal thalamus (which is within their imaging fields of view), then this major discrepancy deserves a substantial discussion.

6) given the maturation of the telencephalon between 6 and 14 days of age, how can the authors justify the brain atlas segmentation in Figure S5I? Anatomical structures present in their fish would not even be present in the 6dpf atlas. They may benefit from manual segmentation or alignment with an adult or juvenile brain atlas (see work from Elena Dreosti, Emre Yakshi, and others). At the very least, the authors should show examples of (or better yet, quantify) their atlas registration.

7) There is no evidence for the neural accumulation idea in the discussion (starting at line 338). I suggest they drop this conjecture, or back it up with evidence that punctate actions from the conspecific drives directed ramping or step-like accumulation to a boundary (movement initiation) in neural populations or single neurons.

8) I am unsure about where the N values are coming from in some cases for statistics. In Fig 3B for example (but also 4B,E and S5), are these cells in an example fish, or pooled across all recorded fish? If the latter, how are the authors handling this nested data organization (non-independent cells within independent subjects) for statistical tests? For decoding accuracy and cosine distances, this is fine because there is one value per independent fish subject. Please clarify either way.

Overall, Lifshitz et al. tackle an interesting topic with new methods, but in many cases are missing details and justification for their conclusions. I think this manuscript could be a good fit for the wide readership of Nature Communications, but it would require substantial revisions and incorporation of the points raised above to be appropriate for publication.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

The authors have addressed and incorporated all of the questions and suggestions in the revised version of the manuscript. At this point, I have nothing further to add or request. The results and the discussion has improved substantially, as specific brain nuclei were contextualized to explain the behavioural changes and the associated neural activity. The additional behavioural experiments (summarized in Figure 5) add value to the work and further support to the conclusions. I am very pleased with the final result.

Reviewer #2

(Remarks to the Author)

My concerns were addressed by the authors with new experiments and analyses. I now recommend acceptance of the manuscript. However, I have two additional minor points that I think should be addressed before formally accepting the manuscript.

1) The authors' characterization of the pallial neurons involved in social interactions has some limitations. Only their rough location in the pallial region was identified, and no information about their molecular identity, neurotransmitter content, and connectivity was provided. Moreover, random neurons in the pallium, and not specifically the ones active during social interactions, were targeted in cell ablation experiments. These limitations should be pointed out in the discussion.

2) I would suggest to add information on number of samples used for the experiments in the figure legends, in addition to or instead of the main text.

Reviewer #3

(Remarks to the Author)

The authors have submitted a substantially improved revision. The added behavioral analyses and clarifications are valuable. The clearer description of conspecific sequences and the more systematic examination of approach versus non-approach movements make the behavioral results more interpretable. The manuscript is clearer and the overall logic of the study is easier to follow.

At the same time, there are significant concerns regarding (i) experimental controls for the ablation experiments, (ii) the internal consistency and interpretation of the neural analyses across brain regions, and (iii) the framing of the methodological and conceptual contributions. I outline these points below.

1. Need for standard controls and experimental detail in the pallial ablation experiment

The inclusion of the pallial ablation experiment is potentially very informative and nicely complements the imaging data. However, in its current form, it does not yet meet standard expectations for zebrafish neural ablation studies. First, the study lacks control (sham) ablations (e.g., applying the same laser protocol in a nearby region or in cells not implicated in approach behavior). For example, the habenula is not implicated in the present study and can potentially serve as a control region. Without such controls, it is not possible to determine whether the behavioral changes arise specifically from the loss of pallial neurons involved in social approach, as opposed to general consequences of the ablation procedure (local tissue damage, inflammation, stress, etc.). Second, the experimental details of the ablation should be more easily accessible directly from the Results. For the reader to evaluate the manipulation, it would be helpful to know approximately how many neurons are ablated per fish and how much residual activity remains in the targeted region (i.e. whether ablation leads to partial versus near-complete silencing). Third, because telencephalic structures lie close to the highly pigmented eyes, which efficiently absorb two-photon excitation light, there is a realistic risk that laser ablations could inadvertently affect the retina or optic pathways. A simple control for visual function would substantially increase confidence in the conclusions; without it, it remains difficult to exclude the possibility that some of the observed behavioral changes reflect unintended visual or other nonspecific effects.

2. Consistency of analyses and interpretation across brain regions

The description of how different brain regions contribute to prediction of approach behavior remains somewhat difficult to reconcile across sections. In the section "Distinct neural processes predict upcoming approach movements on a single-trial level," the emphasis is on midbrain and hindbrain as the regions from which upcoming movement identity (approach versus

non-approach) can be robustly decoded from pre-movement population activity. In the subsequent section “Coordinated brain-wide neural dynamics precede approach behavior,” the narrative shifts to highlight the telencephalon, where neurons are described as showing the expected pattern of elevated activity before approach and reduced activity before non-approach.

If telencephalic neurons show a pronounced and directionally consistent modulation, one might intuitively expect decoding performance in telencephalon not to be substantially lower than in midbrain or anterior hindbrain. However, comparing Fig. 3 and Fig. 4, telencephalon decoding accuracy appears substantially lower than in midbrain and hindbrain. The reason for this discrepancy is unclear from the current text and figures.

3. Use of “spatial factors”, “spatiotemporal integration”, and claims of novelty

The manuscript’s use of terms such as “spatial factors” and “spatiotemporal integration” remains somewhat imprecise, and some of the broader claims about novelty could be more carefully framed.

For instance, the sentence, “The approach toward conspecifics is known to be influenced by various spatial factors, including the number of conspecifics, the visual load they impose on the retina, their shape and their perceived depth”, groups several quantities under “spatial factors.” While perceived depth is clearly spatial, “shape” is more naturally described as a visual stimulus feature, and other relevant dimensions such as color would also fall under “visual” rather than “spatial.” As written, the terminology blurs sensory categories in a way that may confuse readers about what exactly is meant by “spatial.”

Moreover, several prior studies have already examined coordinated trajectories over time. When studies analyze the joint trajectories of two animals, they are by definition considering relationships that unfold across both space and time. In this context, suggesting that previous studies fail to jointly consider spatial and temporal aspects of social behavior is unnecessarily dismissive.

The primary behavioral finding of the present study is that after a conspecific changes direction, the direction of the focal fish’s elicited movement depends on the egocentric position of the conspecific. This is an incremental but valuable refinement of current models of zebrafish social interaction. It would be more effective to state this contribution clearly, without relying on broad or potentially overreaching statements about prior literature.

On methodology:

4. Scope of imaging and use of “brain-wide” terminology

The authors now clearly specify that they image a $705 \times 352 \times 70 \mu\text{m}$ volume located approximately $110\text{--}180 \mu\text{m}$ below the skin. Given that the brain of a 16 dpf zebrafish is around $200 \mu\text{m}$ in depth, this implies that only a fraction of the total brain volume is sampled axially, and that the caudal part of the brain, including much of the hindbrain, lies outside the field of view seen in Supplementary Fig. 1.

In this light, the repeated use of the term “brain-wide” is problematic. Consider the two citations that the authors give in the rebuttal: “Across fish, 10,000-15,000 neurons were identified within this volume per fish, in line with previous brain-wide volumetric studies (Lin et al., 2020; Kappel et al., 2022).”

First, in “Visual Recognition of Social Signals by a Tectothalamic Neural Circuit.” Nature 2022, Kappel et al. use a similar 2P functional imaging approach to the present study. Crucially, Kappel et al. never claim “brain-wide imaging” in the title, abstract, or results. Instead, they carefully describe the exact regions imaged and restrict their conclusions to the network that is actually sampled, explicitly avoiding overextending their methodological claims.

By contrast, Lin et al. 2020 is a true brain-wide imaging study using light-field microscopy, with an imaging volume of $700 \mu\text{m} \times 700 \mu\text{m} \times 200 \mu\text{m}$. This represents an order-of-magnitude more comprehensive sampling of the larval zebrafish brain than in the present work and approaches the common-sense meaning of “brain-wide” by attempting to cover most of the brain, rather than a relatively limited subvolume.

Placing these two papers together under the umbrella of “previous brain-wide volumetric studies” conflates fundamentally different methodological regimes and obscures the care with which each prior study calibrated its claims to its actual coverage. The term “brain-wide” imaging is typically reserved for attempts to sample the brain at least somewhat comprehensively (for example, $700 \mu\text{m} \times 700 \mu\text{m} \times 200 \mu\text{m}$), not for volumes that encompass only a small fraction of it. If such comprehensive coverage were attempted here, the authors could make a much stronger argument that the telencephalon signal they observe that ramp prior to approach movements is genuinely unique in the brain. However, if only roughly 10 % of the brain is sampled, as in the present study, it remains entirely plausible that similarly tuned populations reside in unobserved regions, such as the caudal hindbrain or other omitted areas (e.g., hypothalamus).

The real strength of the present study is that it records activity from several brain regions (telencephalon, midbrain, and an anterior portion of the hindbrain) at cellular resolution during a social assay. That is already substantial. To make the methodological claims more accurate and transparent, I strongly recommend removing the term “brain-wide” from the title and from strong statements in the abstract and discussion, unless the imaged volume is substantially expanded to cover the majority of the brain, and replacing “brain-wide” with more precise descriptions such as “multi-region volumetric imaging”. These changes would not diminish the value of the data; instead, they would increase the credibility of the methodological description.

5. Real conspecific assay versus VR / replayed video stimuli

The authors argue that their assay isolates the social visual cue and preserves the natural 3D kinematics and pose dynamics of a real conspecific, and that this “full visual complexity” is difficult to reproduce in VR. However, VR systems can present not only simplified synthetic stimuli such as dots, but also highly realistic stimuli, including videos or rendered 3D models of real fish. Therefore, contrasting the present assay with VR in terms of “full visual complexity” versus “simplified synthetic stimuli” is misleading, especially when referring to studies that have already used more sophisticated virtual

conspicuous.

Moreover, the current data do not isolate which specific aspects of “full visual complexity” (e.g. 3D kinematics, fine-scale posture, textural details) are needed for the observed differences in neural or behavioral responses.

To present a more balanced and convincing methodological perspective, I suggest explicitly acknowledging that VR-based approaches can replay and systematically manipulate complex, realistic stimuli, including those derived from real conspecific movements, and reframing the comparison in terms of technical ease of implementation: the present assay provides a technically easy method to deliver naturalistic real-conspecific stimuli, while VR enables precise control and causal testing of stimulus components.

Reviewer #4

(Remarks to the Author)

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

Reviewer #5

(Remarks to the Author)

In this revised manuscript, the authors have improved upon their original manuscript, with new experiments, analyses, methodological clarifications, and discussion points. I believe the authors have adequately address the points I raised, and have correctly added context or caveats to the text when appropriate.

In response to the Major points I raised:

- 1) The authors show, in FigS3G, that social approach is correlated between embedded and freely-moving conditions. While this not the strongest correlation, it is nonetheless significant for the low N, and I find this satisfactory. Though, I would prefer the authors not use the word “novel” to describe their assay in either the figures or the main text. Simply call it an “embedded”, or “tethered” assay (for instance, in the x-axis label of Fig S3G, and throughout the text and methods).
- 2) This is a reasonable response – thank you for clarifying.
- 3) This explanation makes sense, but I think it is still a potential issue. Can the authors more explicitly state this caveat (much lower tethered movement rate) in the main text of manuscript, in addition to the methods (516-519)?
- 4) This is much clearer, especially showing the data in Figs 3F, S4D and S5E/F.
- 5) I agree that the loss of mid/hind-brain activity after pallium laser ablation suggests there is a functional link between the forebrain and other brain regions, but there is no clear anatomical connectivity. Therefore, I appreciate the authors’ willingness to remove this speculation from their discussion.

The authors also gave acceptable answers to the minor points, and I appreciated the nuance of the added discussion about thalamic populations, and discrepancy to the Kappel et al. study.

In sum, Lifshitz et al. have improved their manuscript in the revision – addressing my points and seemingly those of the other reviewers. If they can address any remaining points (including my minor additions above), I think this paper would be well suited for publications in Nature Communications.

Version 2:

Reviewer comments:

Reviewer #2

(Remarks to the Author)

The authors address all my concerns in the revised manuscript. I recommend accepting it for publication in Nature Communications.

Reviewer #3

(Remarks to the Author)

The authors have satisfactorily addressed the major comments. The inclusion of sham controls is a particularly valuable and necessary addition, and it substantially strengthens the credibility of the pallial ablation results.

Minor comments:

1. Variations in sample size across figures and analyses are occasionally puzzling. For example: Why does 3G include 18 fish whereas 3H includes 20? Please clarify the exclusion criteria used for each analysis more explicitly in the results section / figure legends.

In the behavioral analyses, the sample size is sometimes reported as $n = 44$, whereas in other cases six animals are excluded, resulting in $n = 38$. While the Methods indicate that these exclusions correspond to animals that did not exhibit measurable social preference, it would improve clarity to explicitly remind the reader of this rationale at the relevant points in

the Results, or to clearly indicate which analyses include or exclude these animals. It would also be helpful to briefly note whether inclusion of these non-social animals qualitatively affects the reported behavioral effects.

2. For p-values that are effectively zero, reporting exact numerical values (e.g. $p = 6 \times 10^{-174}$ in Fig. S3) is not meaningful.

3. Finally, I have some reservations about the framing of the pallial ablation results in the Discussion. The manuscript states that the loss of predictive dynamics suggests a role for pallial activity in the distributed neural signature preceding social approach behavior. While this interpretation is plausible, it may be somewhat overstated. The authors didn't specifically target neurons with approach activity. Thus while the ablating semi-random cells within the telencephalon clearly disrupts telencephalic function, the resulting behavioral and neural effects are not necessarily specific to social approach per se. Alternative explanations, such as impairments in working memory, spatial representation, or other cognitive processes that indirectly affect social behavior, cannot be excluded. I would suggest rewording this section to acknowledge these possibilities and to more explicitly frame the conclusions as reflecting a requirement for intact pallial processing rather than a selective role in social approach alone.

Reviewer #4

(Remarks to the Author)

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

Reviewer #5

(Remarks to the Author)

In this revised manuscript, the authors have addressed the remaining issues I raised in the last round, and found their responses to other reviewers' comments reasonable. I have no further comments.

Open Access This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

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

Revision Outline

Manuscript NCOMMS-25-49635-T, Lifshitz et al.

Overview of changes

We thank the Reviewers for their constructive and valuable feedback. We have addressed all of the reviewers' comments, conducted six additional experiments, added further analyses, and substantially revised our manuscript, figures, and supplementary material. A brief summary of the main changes is as follows:

- **Inter-individual variability:** We added new experiments recording the same individuals under our novel assay and during free social interaction. We show that the variability in social approach captured by our assay matches the variability observed in freely swimming zebrafish, underscoring the ecological validity of our paradigm.
- **Controls for social specificity:** We added control experiments to strengthen the conclusion that the reported brain-wide neural signature is specific to *social* approach and not to general approach movements.
- **Causal role of pallial activity in mediating the behavior:** We added new ablation experiments demonstrating that the identified pallial neurons are not only activated specifically prior to social approach but are also necessary for this behavior. Control experiments confirmed that the observed behavioral changes result specifically from the loss of pallial neurons rather than from experimental procedures. Finally, we included new functional imaging experiments following ablation showing that the telencephalic signature is no longer indicative of upcoming movement, and the distinct brain-wide pattern preceding approach versus non-approach movements in the midbrain and hindbrain also disappears.
- **Validation of turn angle prediction in a social context.** In a new experiment, we recorded social behavior in freely swimming fish and demonstrated that our prediction mechanism, which infers turn angles from tail dynamics, can be trained on different behavioral contexts (hunting, exploration, or social interaction) and still produce similar predictions.
- **Controls for the novel social assay in the dark.** We performed additional experiments under dark conditions (n=7) using the novel assay to verify the light–dark difference in approach behavior.
- **Behavioral and neural analyses:** We performed comprehensive kinematic analyses confirming that approach and non-approach movements do not differ in motor parameters, supporting the conclusion that the neural signature preceding these movements reflects social context rather than motor output. Moreover, we demonstrated that telencephalic, midbrain, and hindbrain neurons that distinguished approach from non-approach movements were not selective for conspecific position.
- **Brain structure segmentation and registration:** We updated the division of brain structures to correct an inaccuracy in the previous version and registered our brains to the new juvenile brain atlas.

Reviewer 1

In this manuscript, the authors investigate the neural basis of social approach behavior at cellular resolution in zebrafish. They report a brain-wide neuronal activation pattern that precedes the movement of the focal fish, characterized by increased activity in a small subset of telencephalic neurons, along with reduced activity in the midbrain and hindbrain. This pattern appears to be specific to social attraction or engagement. Overall, I found the study highly compelling and the results both interesting and relevant. The authors introduce a novel and thoughtfully social assay that effectively captures social interactions with a real conspecific. This approach holds great potential for future studies exploring brain-wide dynamics in diverse social contexts. The manuscript is clearly written, the figures are well-prepared and effectively support the findings, both in the main text and supplementary material, and the experiments are carefully controlled.

We thank the Reviewer for the positive and constructive assessment. We have made the following revisions: (i) Corrected anatomical terminology throughout; (ii) Clarified that our assay preserves the full visual complexity of conspecific cues; (iii) Specified imaging depths and added lateral and top views of the brains imaged; (iv) Performed new analyses showing diencephalic neurons that are tuned to the characteristic movement frequency of the conspecific. These neurons, however, do not predict approach decisions; (v) Localized the pallial population, demonstrated their necessity via ablation, and proposed potential encoding of the social cue's reward value. We address each point below.

Major

Throughout the manuscript, the telencephalon is referred to as the forebrain. While the telencephalon is indeed a part of the forebrain, along with the diencephalon, the analyses of these two brain regions were performed separately, as shown in Figure 3A. Therefore, it is inaccurate to refer to the telencephalon as the forebrain. This should be corrected throughout the text: the term "forebrain" should be replaced with "telencephalon," as it is the telencephalon the authors are actually referring to.

We thank the Reviewer for this important correction. We revised the manuscript to use telencephalon (rather than forebrain) where appropriate and updated Figure 3A accordingly.

Additionally, in lines 329–330, the text states: "These forebrain neurons were primarily located in the pallium (dorsal forebrain, Figure S5I)." However, the pallium and subpallium are subdivisions of the telencephalon, not of the forebrain. This should also be corrected for anatomical accuracy.

We corrected the text to state that the pallium and subpallium are subdivisions of the telencephalon, improving anatomical accuracy (lines 356-360).

lines 310–312: In this section of the discussion, the authors state that the full complexity of social cues was preserved. However, based on the description provided in the Methods section, it appears that only visual cues were accessible during the social assay. The fish were placed in separate compartments, preventing the exchange of chemical or mechanosensory cues in the water. This point should be clarified and the text revised accordingly to accurately reflect the sensory modalities available during the assay.

We have clarified this point. Our assay preserves the full complexity of *visual* social cues from a real conspecific, including naturalistic motion, depth, and shape, which have been shown to drive substantially stronger responses than synthetic or 2D visual stimuli (Larsch & Baier, 2018; Kappel et al., 2022). We have revised the text to specify "visual social cues" (lines 333-335).

Perhaps I missed this information in the Methods section, but what was the imaging depth during data acquisition? Was the entire brain imaged from dorsal to ventral? I ask this in relation to the acquisition of the neuronal activity in the diencephalon. In Figure 3A, the diencephalon is depicted as a very small grey area that does not represent the size or entire location of this brain region. Anatomically, the diencephalon lies beneath the optic tectum and is a considerably large structure. Was the diencephalon fully captured? This point is unclear in the manuscript, and the representation in Figure 3A is somewhat misleading. A lateral view of the brain would provide a more informative depiction of the imaged areas.

We thank the Reviewer for this important point. Our imaging depth ranged from 110–180 μm below the skin surface, capturing the dorsal diencephalon. This configuration provided broad brain coverage while maintaining sufficient temporal resolution. We have now (i) specified these depths in the Methods (lines 437–442), (ii) added top and lateral views of the imaged volume (Figure S1B), and (iii) revised Figure 3A for anatomical accuracy. These additions clarify which regions were sampled and provide a complete account of our imaging volume.

If the diencephalon was fully acquired and does not show any distinct activation pattern, it would be very interesting to include this topic in the discussion and propose potential hypotheses as to why this brain region, typically involved in the expression of social behavior and considered part of the social behavior network, does not respond in a social affiliation assay.

We thank the Reviewer for this comment. We performed new analyses showing that dorsal thalamic neurons respond more strongly when the conspecific moves at its characteristic bout frequency (Figure S6D), consistent with prior work (Kappel et al., 2022). However, despite this frequency-tuned encoding of the visual social cue, dorsal thalamic activity did not predict upcoming approach vs. non-approach movements (Figure S6E). Additionally, we added a paragraph to the Discussion summarizing the functional profile of the dorsal diencephalon in our assay and noting that ventral diencephalic nuclei, which were not sampled here, may exhibit action-specific patterns (lines 370–379).

Still in the discussion, it would be interesting to go beyond description and explore why more dorsal areas of the telencephalon are active. Which specific brain nuclei are located in these regions? Could this activity be related to the reward system?

We have substantially expanded the analysis of the dorsal telencephalon. Spatial characterization showed that these neurons are located in the pallium (Figure S6A). Targeted ablation of this population abolished approach behavior, demonstrating its necessity (Figure 5D–G). In addition to changes in behavior, the brain-wide predictive signature also disappeared following pallial neurons ablation (Figure 5H,I). In the Discussion, we now note that the pallium is the teleost homolog of limbic-related circuits, suggesting that its activity may reflect valuation or reward assignment (lines 359–360).

Minor

Line 395: The text states, “All details of the control experiment (Figure S6G,H),” but there is no Figure S6. The last figure in the supplementary information is Figure S5. This should be corrected.

Corrected. We now reference to Figure 5A–C (moved to main text).

lines 338 and 346: Both paragraphs begin with similar phrases, “Our results uncovered” and “Our data revealed”, which is somewhat repetitive. Consider rephrasing one of them.

Revised to eliminate repetition.

Figure panels: In some figures, such as Figure 1I and others, the panels are divided into top, middle, and bottom sections. For clarity, it would be helpful to label these subpanels directly in the figure as I, I', and I'', for example, or identify them in some other way.

Implemented, as suggested by the Reviewer, throughout the manuscript. We added i, ii, iii for cases with independent subpanels and non-shared axes; otherwise, we maintained a single axis scheme for clarity.

Caption of Figure G: The caption states, “. . . showing that approach vs. non-approach distinctions emerged earlier for long sequences.” However, the comparison here is not approach vs. non-approach, but rather is approach and not approach. This should be revised.

To clarify, the comparison in Figure S4H (updated version) is between approach and non-approach movements, consistent with the definitions used throughout the paper.

Reviewer 2

In this paper, Lifshitz et al., describe a novel assay for monitoring at the same time behavioral responses and neuronal activity, via calcium imaging, of a zebrafish larva in the presence of a conspecific. They used this setup to identify patterns of neuronal activity correlating with approaching of the imaged fish toward a freely swimming conspecific larva. The assay is novel and potentially very useful for studying the neuronal mechanisms underlying social behavior in zebrafish. Moreover, I think that the experiments were well designed and mostly support the authors' conclusions. However, I find the work extremely correlational and lacking any supporting evidence that the observed patterns of neuronal activity, or any of the neurons active during approach behavior, are actually required for social interactions. It is my opinion that in the current state the manuscript does not provide new insights into the neuronal mechanisms of social behavior, and should not be accepted for publication before a major revision. Please, find below my specific comments.

We thank the Reviewer for the constructive feedback and for recognizing the novelty and utility of our assay. We have made the following revisions: (i) We carefully addressed the Reviewer's main concern that the work was correlational. In a new set of experiments ($n=9$), targeted laser ablation of the identified pallial neurons significantly reduced social approach while leaving movement kinematics intact; (ii) In another experiment, we show that following pallial ablation, the brain-wide signature preceding approach behavior was abolished; (iii) We added recordings of freely swimming fish during social interaction to demonstrate that our tail-movement readout is robust across different training contexts; (iv) To show social specificity, we added more controls and performed the requested two-sided statistical test; (v) As suggested by the Reviewer, we moved the neural results on social specificity to the main text and included the additional ablation results. Together, these revisions establish both specificity and causality, as the Reviewer requested. We respond point-by-point below.

Major

Why did the authors used prey capture behavior in freely moving fish to compare turning angles in head-fixed larvae to the head-fixed preparation? Wouldn't it be more appropriate to use turning toward conspecific in freely moving fish for the comparison and validation of the tail movement quantification method? Are not turning parameters expected to be different in the two contexts (i.e., J-turns during hunting?)

We thank the Reviewer for this important point. Our original analysis included movements during both hunting and spontaneous exploration, which yielded similar predictive features. To directly address the Reviewer's concern, we have now added movements from freely socially interacting fish. We show that the principal components describing tail movements (Figure S2E), as well as the resulting turn angle predictions (Figure S2I), remained consistent, regardless of the behavioral context on which the model was trained (hunting, exploration, and social interaction). The text has been revised accordingly in the caption of Figure S2, as well as the Methods section (lines 491-492, 503-511).

The authors stated that there is an effect of position of conspecific fish on turning angle of the focal fish (figure 1G), and that the effect is not present in the dark (Figure 1H). The effect is quite small and the authors did not perform statistical test to support these two conclusions. This should be done.

We thank the Reviewer for this helpful comment. We have now performed the appropriate statistical analyses. We found a significant effect of conspecific position on the approach probability of focal fish in the light (Figure 1G; binomial test, $p = 2 \times 10^{-7}$). Notably, the strength of this effect matches the effect observed in freely interacting fish (Harpaz et al., 2021). Since relatively few movements were captured in the dark during the first 5 minutes of each experiment, we expanded the dark dataset by recording fish kept in darkness throughout the experiment ($n=7$), repeated the analysis, and confirmed that this effect was absent in the dark (Figure 1H; binomial test, $p = 0.97$). These results have been added to the caption of Figures 1G,H.

The authors claims that distinct neural activity preceding approach vs. non-approach movements emerged only when focal fish observed a real conspecific (lines 298-299). The difference is small, suggesting that the neuronal activity pattern observed may mostly underlie activity related to a generalized approach response, rather than one specific to social interactions. The fact that activity of only a small group of neurons is responsible for the difference between generalized and social-specific approach could be used to identify and better characterize these neurons.

However, this detailed analysis was not performed and, as a consequence, the manuscript lacks any data related to identity of the neurons responsible for social interactions. More effort should be put to better characterize the difference between neuronal activity correlating specifically with social interactions. Moreover, I'm not convinced that there is an actual statistically significant difference in the midbrain activity in Figure S5G. Why did the authors perform one-sided t-test instead of a two-sided one? Finally, I think that data supporting the presence of neuronal activity correlated specifically with social approach is a very important piece of evidence for this manuscript, which should not be relegated to a supplementary figure.

We thank the Reviewer for raising these important points, which prompted us to strengthen the characterization of these neurons, their specificity and their causal role in social behavior.

Social specificity. To compare cosine distance dynamics when fish were presented with a real conspecific versus a synthetic projected dot, we initially used a one-sided t-test because our hypothesis was directional and based on a consistent effect in the larger main dataset. Nonetheless, a significant main effect of stimulus type (conspecific vs. moving dot) was present even before the post-hoc comparisons (see caption of Figure 5B). To directly address the Reviewer's concern, we performed an additional control experiment and re-analyzed the data using a two-sided t-test, which also yielded a significant effect (midbrain conspecific: $p = 0.026$, hindbrain conspecific: $p = 0.005$). This confirms that midbrain and hindbrain activity differ between approach and non-approach movements only in the presence of a real conspecific (Figure 5B). Moreover, we now show that across fish, the distinct telencephalic and midbrain neural activity preceding approach and non-approach movements was present when fish observed a real conspecific (Figure 5C), but absent when they viewed a moving dot, further validating the social specificity of these neural dynamics.

Neuronal characterization. We characterized these cells anatomically and functionally as follows: (i) We now identify these neurons as pallial neurons (Figure S6A); (ii) We show that they are necessary for social behavior using an ablation experiment (Figure 5D-G); (iii) In an additional experiment, we performed functional imaging under our novel assay following ablation and found that not only was the pallial signature abolished, but also the midbrain and hindbrain signatures disappeared (Figure 5H,I); (iv) We carried out new analyses of the functional properties of these neurons and found that their activity cannot be explained by the position of the conspecific (Figure S5B,C), indicating that they do not encode low-level sensory features of the social cue. Together, these new results further characterize these pallial neurons, establish their causal role in driving social behavior, their specificity to social behavior, and their anatomical position.

Figure placement. As suggested by the Reviewer, we moved the social specificity results from the Supplementary Material to the Main Figures (Figure 5A-C) to highlight their importance and better integrate them into the manuscript's narrative.

The major shortcoming of this work is that it is only a correlational study. There is no evidence that the neuronal activity pattern observed or any of the neurons active during an approach are required for social behavior. The authors should perform studies to test the effect of ablation or optogenetic perturbation of the identified neurons on approach toward conspecific fish. The following are very good examples of this type of approach: Vladimirov et al. Nat. Methods (2018) <https://doi.org/10.1038/s41592-018-0221-x>; Kappel et al. Nature (2022) doi: 10.1038/s41586-022-04925-5

We thank the Reviewer for this critical point, which motivated us to extend our study beyond correlation. Following the approach outlined in the studies suggested by the Reviewer (Kappel et al., 2022), we performed a new ablation experiment ($n=9$). These experiments included a sequence of behavioral screening before ablation (Dreosti et al., 2015), targeted laser ablation of the identified pallial neurons, 24 hours of recovery, and post-ablation behavioral screening. Ablation significantly reduced social preference (Figure 5F,G), while leaving kinematic features intact (Figure S6B), demonstrating that these neurons are not only active prior to social approach but also causally required for social interaction. In a subsequent set of experiments, we show that ablation of pallial neurons also disrupts the midbrain and hindbrain signature preceding approach behavior (Figure 5H,I). These new experiments are detailed in the new Results section (lines 307-328) and in the Methods section (lines 616-633).

Minor

It's not clear what brain regions the authors consider to be part of the forebrain in their Ca²⁺ imaging analysis, since they separated forebrain and diencephalon, which is part of the forebrain. Did they mean telencephalon instead of forebrain? If so, the term should be corrected.

We thank the Reviewer for pointing this out. Indeed, we meant telencephalon rather than forebrain, and we have corrected this terminology consistently throughout the manuscript.

Figure 2E. Why are most of the points representing conspecific angular position clustered between -100 and 0 degrees? The distribution in Figure 1E doesn't look that skewed.

We thank the reviewer for this observation. While Figure 1E shows the distribution of conspecific angular position across all recorded fish, Figure 2E illustrates an example from a specific pair in which the conspecific spent more time on the left side of the focal fish. Such skewness is evident at the level of individual fish but averages out across the population. The tendency of locked movements to be directed toward the conspecific at the fish population level is demonstrated in Figure 2F.

Reviewer 3

The authors have implemented a social interaction setup where a “focal” fish is head restrained and tail free, and a freely moving conspecific is placed at a distance (separated by a barrier). Using this setup, the authors have reproduced previously published observations that fish will preferentially reorient in the direction of conspecific motion. By recording from a fraction (10,000 neurons) of the “focal” fish brain at a time, the authors find neural activity patterns that are correlated with movement toward conspecifics. While I found the subject of the manuscript to be interesting, there are significant issues with data interpretation and the overall validity of the conclusions. The advantage of this experimental setup over previous publications is unclear. The Methods section lacks essential details.

We thank the Reviewer for the thoughtful critique. In the revision, we (i) clarified the imaging protocol and anatomical registration, adding schematics and complete parameter reporting; (ii) strengthened analyses to rule out kinematic confounds and angular position tuning, thereby reinforcing our interpretation; (iii) defined and illustrated “conspecific sequences”, including their spatial distribution; and (iv) added a manipulation, targeted laser ablation of pallial neurons, that reduced social preference within individuals. Together, these additions improve methodological transparency, strengthen data interpretation, and provide causal validation of our conclusions.

On methodology:

The advantages, as well as the implementation of the experimental setup, are both unclear. The experimental setup is not sufficiently described, particularly regarding the physical properties of the barrier separating the fish. Critical details such as the material's structure, thickness, refractive index, and how it might affect visual perception at various angles are missing.

All hardware details (materials, dimensions, CAD files) and catalog numbers are now provided. Briefly, we used standard optically transparent petri dishes for the inner and outer chambers (Nunc 171099: 35x10 mm, Corning Gosselin: 55x15 mm, thickness: 0.6 mm, refractive index: ≈ 1.59). These details are now added to the Methods section (lines 418-422).

What is the advantage of this system? This is arguably a more simplistic and less sophisticated setup than a VR system. VR for head-restrained larval zebrafish has been implemented in multiple labs, including closed loop VR. Given that the authors have deliberately removed olfactory or somatosensory information, what can this offer that a visual display cannot?

Our assay isolates the social visual cue and presents its natural statistics, including a real conspecific's 3D kinematics and pose dynamics, features that remain difficult to reproduce faithfully in VR. The live stimulus carries ethologically relevant features (such as fine postural changes, acceleration transients, self-occlusion, natural variability), enabling the discovery of neural signatures of social approach under realistic visual input while excluding non-visual modalities. Consistent with this, prior work reports weaker social attraction to simplified synthetic cues (under VR conditions) compared to real conspecifics (Larsch & Baier, 2018; Kappel et al., 2022). Additionally, the setup is simple, robust, high-throughput, requires minimal calibration, avoids display or refresh artifacts, and is fully compatible with long, stable two-photon imaging sessions. This assay and VR paradigms are complementary: the real-conspecific assay identifies neural responses under natural visual statistics, and VR can then test causality by isolating candidate features under precise control. We now discuss this explicitly (lines 333-335).

The description of how neural activity was recorded is vague and lacks important technical details. “Subsequent to the functional imaging, the entire brain structure (300-350 imaging planes, 1 μm apart) of several focal fish ($n=6$) was scanned for structure.” – What does “scanned for structure” mean? Is 300-350 imaging plane collected per fish, or summed across all fish? At what temporal resolution? For how long? For how many sessions per fish?

We appreciate this and have clarified the Methods. “Scanned for structure” refers to a post-hoc anatomical stack acquired at the end of each functional session (no time series), used for registration and anatomical segmentation. After two-photon functional imaging (30 min total), we collected 300–350 z-planes at 1 μm steps per fish (not pooled across fish), covering the entire brain, including the volume imaged for function. Functional time-series data were acquired volumetrically at 3 volumes/s for 30 min (5 min dark, 25 min with a visible conspecific). Each fish underwent one session, unless noted for control experiments. All imaging parameters are now reported explicitly

in the Methods (lines 437-451) and illustrated in Figure S1A,B.

2b. Previous to this statement, the authors write “Simultaneously, volumetric two-photon imaging (Nikon objective, CFI $\times 16$, 0.8 NA) captured neural activity over a $705 \times 352 \times 70 \mu m$ area at 3 volumes/sec. This is clearly not the entire brain volume. So then what did the authors actually choose to image for a given animal? How many biological replicates exist for each brain region?

We imaged a $705 \times 352 \times 70 \mu m$ volume located 110–180 μm below the skin surface. Top- and side-views of the recorded neurons are provided (Figure S1B), and the dorsal-ventral boundaries are specified (lines 437-442). This configuration trades volumetric coverage for higher temporal resolution. The same volume was acquired repeatedly as a multi-plane stack covering all regions within that volume (i.e., not scanned region-by-region). Across fish, 10,000-15,000 neurons were identified within this volume per fish, in line with previous brain-wide volumetric studies (Lin et al., 2020; Kappel et al., 2022).

2c. In the Results section, the author separately state that they “record activity from over 10,000 neurons across the brain of the focal fish” – this is only 10% of the larval brain, which region was recorded from each fish? How long was the recording? How many sessions were recorded per fish?

We thank the Reviewer for this important point. Within the imaged volume, including the dorsal telencephalon, dorsal thalamus, tectum, and hindbrain, Suite2p identified over 10,000 neurons per fish (iscell parameter > 0.5). This number is consistent with previous functional volumetric recordings in zebrafish, as mentioned above. We have now substantially revised the Methods section and added supplementary panels (Figure S1A,B) to provide a detailed description of the experimental assay and imaging protocol (lines 437-443). Briefly, each fish was recorded for 30 minutes: 5 minutes in darkness followed by 25 minutes with a visible conspecific. Volumetric imaging captured all regions simultaneously during each 30-min session.

3. 16 dpf brains are registered to a 6 dpf reference brain? Given the size and structural difference between the brains at these developmental ages, this can only give crude anatomical annotations at best.

We have corrected and clarified this point. All 16 dpf brains were manually annotated, and a subset of fish ($n=6$) was registered to a 16 dpf reference brain (one of the imaged fish). We used an atlas only to localize telencephalic neurons within the pallium. This is now done using a 21 dpf atlas (Brain atlas of juvenile zebrafish currently available online from the Dreosti lab), which shows that these neurons fall within pallial boundaries (Figure S6A). We have updated the Methods accordingly (lines 448-458).

On data analysis and interpretation:

1. The authors conclude based on the topographical organization of activity in the tectum that “This confirms that the focal fish observed the natural social cue and encoded its physical properties, such as position and direction of motion, in tectal neural activity.” This is a somewhat trivial statement. The retinotopic organization of activity in the optic tectum has long been established. Any object (moving or not moving, social or non-social) will be encoded topographically in the tectum.

We agree that tectal retinotopy is well established. Here, the value is validation: it confirms that the focal fish viewed a freely moving conspecific and that tectal activity tracked the cue’s position and direction under our naturalistic social paradigm. To our knowledge, this is the first demonstration of retinotopic tectal responses to a freely moving conspecific (as Reviewer 5 noted, “the tectal map is a nice validation of social vision”). In response to Reviewer 5’s request, we have now added an additional validation, showing tectal direction selectivity map (Figure S1F). We include these analyses solely to confirm visual engagement and retinotopic integrity under our configuration.

More importantly, the authors have not demonstrated that any of the behavioral or neural results are specifically social (see 3 below).

Please see our response to Point 3 below. We added control experiments and a comparison demonstrating social specificity of the neural signature.

2. The statement in the abstract “We demonstrate that zebrafish approach behavior is jointly shaped by spatial and temporal information from conspecifics rather than by these features acting independently” is also expected given

existing publications. It's known that zebrafish respond to the joint position and motion information of conspecifics. Zebrafish also responds jointly to the position and motion of many other stimuli (e.g. prey, predators, objects, grating, etc.).

We thank the Reviewer for this comment. While spatiotemporal integration is often assumed, previous works in the social context of larval zebrafish demonstrated the effects of position (Hinz et al., 2017; Harpaz et al., 2021) and timing (Dreosti et al., 2015; Larsch & Baier, 2018; Stednitz & Washbourne, 2020; Stednitz et al., 2025) on shoaling behavior separately. The novelty of our findings lies in demonstrating their interaction: approach movements of the focal fish are more likely to be temporally coupled with those of conspecifics (Figure 2E-G). Moreover, we show that the approach probability and synchrony level of the focal fish are related and account for individual differences in social behavior (Figure 2H). Notably, recent unpublished work addressing a similar question (spatio-temporal interaction in the social context) in Medaka fish (Harpaz et al., 2025, bioRxiv) underscores the timeliness and relevance of our findings. We have clarified the corresponding sentence in the abstract (lines 12-14).

The authors claim that the “time lag between fish movements was significantly shorter than expected by chance”, but the right control is not chance but stimulus motion that do not resemble fish movement, which has been demonstrated in previous studies (e.g. Kappel et al 2022), but not in the present study.

We thank the Reviewer for the opportunity to clarify. The goal of our analysis was to test whether focal fish synchronize their movements with conspecifics more than expected by random coupling. For this question, the appropriate control is a shuffled (chance-level) time-lag distribution, which we used to establish whether the observed synchrony was statistically shorter than chance. By contrast, Kappel et al. (2022) addressed a different question: whether fish preferentially respond to bout-like movements over arbitrary, non-fish-like motion. Their comparison to synthetic stimuli that do not move like fish was, therefore, essential for testing stimulus selectivity, whereas our shuffled control is appropriate for testing temporal coordination. Thus, our analysis and theirs address complementary but distinct hypotheses, each requiring different controls.

One of the major issues in this manuscript is the vague definition of “approach movements,” which are classified as any movement toward the conspecific. This definition creates a potential bias, as movements toward a conspecific may be statistically different from those directed elsewhere. Without careful characterization of the kinematic parameters (e.g., bout speed, inter-bout interval, duration, type), it is difficult to disentangle whether observed neural differences are a result of movement patterns, sensory representation, or social approach. The lack of detailed analysis in this regard significantly weakens the authors' conclusions about the neural basis of social approach. For example, the authors only present a single bar plot suggesting the “predicted tail angle” is not different between approach or non-approach movements. Imagine a very simple scenario, where the IBI of consecutive “approach movements” is significantly lower than the IBI of consecutive “non-approach movements”. Differences in movement rate correlate with differences in global brain activity, something that has been demonstrated across species. Without taking this into account, the authors would then misinterpret any observed difference as reflective of social processing.

We appreciate the Reviewer's concern regarding the definition of approach movements and the possibility that neural differences might reflect generic kinematic features rather than social approach. To directly address this, we performed new detailed analyses of multiple movement parameters. Specifically, we quantified predicted turn angle, inter-bout-interval, movement duration, tail amplitude, and tail frequency for approach versus non-approach movements and found no significant differences across any of these parameters (Figure S4F). These analyses demonstrate that approach movements are not simply a subset of faster, shorter, or more frequent movements, and therefore cannot be explained by differences in bout kinematics or movement rate. Consequently, the neural differences we report reflect social approach itself, rather than generic motor activity. We have added this finding to the Results section (lines 212-215).

5. A number of conclusions are drawn from the so-call “conspecific sequences” (e.g. the “focal” fish is claimed to respond preferentially to the early phase of these sequences). What exactly are these sequences? Given the geometric constraints of the environment and movement patterns of zebrafish, one could imagine that these are mostly just the fish turning around at each end of the chamber. During the “early phase” of these leftward or rightward sequences, the conspecific is simply positioned closer to the optimal response region of the “focal” fish. Furthermore, the conspecific seems to have a spatial preference for a particular zone of the outer chamber, which

further complicates the analysis. It's unclear what is the cause for this, and whether it is related to the presence of the "focal" fish at all.

We thank the Reviewer for pointing this out. Apparently, the point regarding the conspecific sequences was not clearly delivered. We now define and illustrate conspecific sequences (Figure S3Hi) and show that the distribution of angular positions at which conspecifics initiate a sequence covers the entire range of the chamber (Figure S3Hii). Sequences can start anywhere in the outer dish, so the early phase can occur at many locations (in fact, it occurs less frequently at the edges of the chamber). Thus, early focal responses cannot be attributed to a fixed positional bias. Regarding the preference of conspecifics to spend more time in the middle of the outer plate, this tendency may reflect the spread of projected light across the curved outer plate, which creates a slight illumination gradient (brighter at the center and dimmer toward the edges), drawing conspecifics toward the middle. In this study, we did not examine whether the conspecific actively observes the focal fish, as it does not serve as a natural social cue. Accordingly, our analyses focus solely on the behavioral and neural responses of the focal fish to the conspecific.

6. The authors go on to state that: "our results indicate that social movements in zebrafish arise from long-timescale neural computations, linked to the time elapsed between a change in conspecific direction of motion and focal fish movement onset." There is no clear support for this conclusion. Multiple previous publications, e.g. Stednitz et al 2018 or Dunlap et al 2021, have already demonstrated that the forebrain responds to conspecifics. Thus, given that approach movements are correlated with angular position of the conspecific, the simplest explanation for much of data is that forebrain activity is modulated by the position of conspecific. Imagine that a forebrain neuron is simply tuned to the angular position of the conspecific, peaking around 50 degrees, where the fish is most likely to turn towards the conspecific. Given the movement statistics of larval zebrafish, a conspecific is likely to spend several seconds (at least) within this angular range on each traversal through the environment. In this case, even if induction of turning towards to the conspecific is purely probabilistic (i.e. having nothing to do with evidence accumulation), when averaged over many movements, one would expect elevated forebrain activity for several seconds preceding an "approach movement".

We thank the Reviewer for raising this important alternative explanation. We tested whether telencephalic activity preceding approach movements simply reflects angular tuning to the conspecific position. We analyzed the activity of neurons identified with our "approach regressor" and found that these neurons showed elevated activity before approach movements and reduced activity before non-approach movements (Figure 4D,E). This distinction was consistent across angular positions of the conspecific, rather than being biased toward any particular position (Figure S5B). Importantly, neither these neurons nor the midbrain or hindbrain neurons with the opposite pattern encoded the angular position of the conspecific (Figure S5C). Thus, the elevated telencephalic activity cannot be explained by angular tuning, but instead reflects an approach-specific computation. This has now been added to the main text (lines 270-274). Our conclusion regarding the relation between conspecific sequences and the timing of the neural distinction is based on the finding that a longer duration between the onset of a conspecific sequence and the onset of a focal fish movement was associated with a longer neural distinction preceding movement onset (Figure 3F, Figure S4H).

7. Other possible explanations for persistent activity differences before a given movement, such as the statistics of the preceding set of movements or arousal, are not considered at all.

We thank the Reviewer for raising this important point. Inspection of signal preceding and following single movements (Figure S4D, Figure S5E) shows that this persistent activity decays after each focal fish movement and ramps up again prior to the next movement. To eliminate potential contributions from preceding movements, only movements that were not preceded by another movement within the selected time window were included when averaging across movements (lines 590-591). Additionally, we agree that factors such as global arousal state may contribute to persistent activity differences observed before social approach. While we did not measure arousal-related variables (e.g., heart rate, pupil diameter) in the present study, we acknowledge a potential relationship between the observed neural dynamics and internal states (lines 388-390), and have added a reference specifically addressing the effects of arousal (Lovett-Barron et al., 2017).

8. Throughout, the authors have attempted to correlate neural activity to behavior. To claim that the observed activity patterns "mediate" or drive social approach would require functional manipulations, which have not been

performed.

We thank the Reviewer for this important point, which motivated us to extend our study and perform functional manipulations. We added a new experiment (n=9) using targeted laser ablation of the identified pallial neurons (lines 616-633). Social preference (Dreosti et al., 2015) was measured within subjects pre- and 24 h post-ablation. Social preference decreased significantly post-ablation (Figure 5F,G) while basic kinematic features remained unchanged (Figure S6B), indicating that these pallial neurons are necessary for social approach. Moreover, in an additional experiment (n=5), we show that following pallial ablation, the brain-wide neural signature preceding approach behavior disappeared (Figure 5H,I), emphasizing the critical role of this region in the neural processes underlying social approach. These experiments have been added to a new section of the Results (lines 307-327).

Altogether, a number of basic analyses and controls are missing. The presentation is confusing, several key interpretations are questionable, and experimental details are missing. This is not to say that valuable insight cannot be gained from the data that the authors have produced with more in depth and carefully considered analyses in the future.

We thank the Reviewer for the comments. We believe that the above additions improve methodological transparency, strengthen data interpretation, and provide causal validation of our conclusions.

Reviewer 4

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

We thank Reviewer 4 for co-reviewing this manuscript and for their contribution to the evaluation process.

Reviewer 5

In this manuscript, Lifshitz et al describe a behavioral assay for late-stage zebrafish larvae, where a head-fixed fish is able to respond to a conspecific swimming in an enclosure within their field of view. This assay allows for brain-wide functional imaging, and the authors describe a multi-regional activity pattern that is predictive of the movement direction in the fish observing conspecific swimming. This is a novel behavioral assay for head-fixed larval zebrafish, and they quantify some interesting neural activity patterns of the observer fish in response to the position/motion of the conspecific swimming in an enclosure around the animal. This paper is tackling an interesting topic, takes a careful approach to distinguish it from previous work on this topic, and is written in a clear fashion. However, I have some concerns about the methods, data presentation, interpretation of results, and integration with the published literature. I detail these points below, and encourage the authors to address these points in a substantially revised manuscript.

We thank the Reviewer for the thoughtful critique. In the revision, we (i) validated the variance in head-fixed, tail-free paradigm against the variance in an established freely moving assay within the same individuals, showing they are correlated; (ii) clarified tail-angle sign alignment across imaging geometries; (iii) explicitly addressed rate differences between head-fixed and freely moving preparations and used empirically observed rates in our Poisson modeling; (iv) extended the analysis window to reveal predictive dynamics over longer timescales; (v) added ablation experiments to examine the causal role of the pallium in social approach; and (vi) better positioned the work within the current literature. These changes sharpen the methods, strengthen the analyses and data presentation, and enhance the interpretability of the results.

Major points

1) The authors state that 2-week old fish approach conspecifics, but this is debatable. While some groups claim there is social this approach by 2 weeks (refs 16,19), others find this attraction only emerges at 3 weeks (for instance, Larsch & Baier, 2018). This makes the choice of 2 week fish problematic, but I understand why it is technically more challenging to image older zebrafish. That being said, the authors appropriately discuss how this behavior is variable across fish and not very robust at this age (30% do not show it here). To validate this assay better in these animals, especially in light of the variability amongst individuals, I would like to see a comparison of the head-fixed and freely-moving social attraction in individual animals. How does the distance between fish in the freely-moving vs head-fixed cases (and subtended visual angle of those conspecifics) compare when looking at approach movements? This assay would be better justified if the authors could show that the most social of 14dpf animals in the free-moving condition are also the most social in the head-fixed condition.

We thank the Reviewer for raising this important point. We agree that social attraction is variable at 2 weeks. To address this, we measured Social Preference Index (SPI) of freely moving individuals and then tested the same fish on our novel assay. SPI in the freely moving assay correlated with approach probability in the head-fixed assay ($n=11$, Figure S3G), indicating that our novel assay not only preserves the variance of the cohort but also faithfully reports each animal's social level. These results provide strong justification for using 16 dpf larvae despite the known variability at this stage. We added these results to the main text (lines 156-159) and the Methods section (lines 539-549).

Furthermore, in the head-fixed prep, does the direction of movement for the focal fish and the conspecific match? Or, is it just that the focal fish moves in the direction of the side where the conspecific is, regardless of their movement direction?

The focal fish's movement direction is not aligned with the conspecific's instantaneous movement direction. Rather, focal movements are biased toward the position of the conspecific, independent of its heading direction (Figure S1D). We now state this explicitly in the caption of Figure S1D.

2) The tail tracking algorithm is quite nice, and has many applications – probably best for hunting assays (Fig S2). But, the behavioral tracking model appears to be trained from freely-moving fish (filming fish from the top), yet applied to head-fixed fish (filming fish from the bottom of the dish). Therefore, the authors should apply a sign-inversion to tail angles recorded from the bottom, where a left-turn from below would be a right-turn from above. Without this correction, the slight attraction reported here (Fig 1G) could in fact be a slight avoidance. This avoidance would be consistent with previous work in 2-week zebrafish (Larsch & Baier, 2018; Harpaz et al., 2021), but would have a very different interpretation from the framing of this manuscript. Please clarify whether this was done in the methods.

We aligned the principal components of tail movements across all datasets - freely moving (exploration, hunting, social) and head-fixed, so that left-right polarity is consistent. This is now stated explicitly in the Methods (lines 508-509). Importantly, even if a left-right inversion occurred, it would apply equally to both the focal fish and the conspecific, ensuring that approach movements still correspond to turns toward the conspecific. As requested by Reviewer 2, we additionally include PCs derived from socially interacting fish, demonstrating that these PCs are robust across social contexts (Figure S2E).

3) The head-fixed assay is clearly useful for the neural access it provides, but it causes fish to behave in an unusual manner that was not discussed enough. For instance, are there statistical considerations for the > 10 -fold difference in movement rates between the head-fixed observer (< 0.1 Hz) and freely-moving fish (> 1 Hz)? Furthermore, is the freely-swimming fish showing any restraint stress? How long do the animals habituate before the behavior/imaging? Finally, is the Poisson model applied here (Fig S3) using these disparate movement rates? This would be important to model appropriately.

As shown in Supplementary Figure S1C, head-fixed fish exhibited movement rates approximately 10-fold lower than freely swimming fish, consistent with previous studies in behaving head-fixed fish (Lin et al., 2020; Wee et al., 2022). We account for this difference by aligning all behavioral and neural responses to the movement onset of the focal fish. Freely swimming conspecifics showed no signs of restraint stress and maintained a stable movement frequency comparable to home tank conditions. Focal fish were allowed to habituate in the dark for 30 minutes prior to imaging to minimize handling effects, and this information is now added to the Methods (lines 427-428). Finally, the Poisson model (Figure S3C,D) was applied using empirically observed movement rates of each fish. We clarify this explicitly in the Methods (lines 516-519).

4) A key conclusion of this paper is about the long neural timescales seen. Cosine distances and neural activities are clearly different for the entire 10s pre-stimulus period in most plots. This is quite interesting, but difficult to interpret without knowing when these changes began. Can the authors plot a longer pre-stimulus period to find out when neural activity diverges between the type of upcoming movements? This is a key question that is central to the authors claims, but not examined sufficiently. If there are "long timescales" of activity predicting social engagement, what is that timescale? It is clearly longer than 10s. Also, this type of > 10 s mid-/hind-brain activity predictive of upcoming movement direction in larval zebrafish has also been observed in a non-social behavior (Lin. . . Vaziri, Cell 2020), which was not cited.

We thank the Reviewer for highlighting this key point. In the revised manuscript, we have extended the analysis window beyond 10 seconds prior to movement onset (Figure S5F) and now show, on a movement-by-movement basis, when neural activity begins to diverge between approach and non-approach movements (Figure S5E). This analysis demonstrates that divergence consistently emerges seconds prior to movement onset and returns to baseline afterward. We also show that the ramping timescale depends on the dynamics of the conspecific: shorter conspecific sequences are associated with shorter ramp durations (Figure 3F, Figure S4H). Finally, we agree that these results are consistent with previous reports of long-timescale predictive activity in non-social behaviors (Lin et al., 2020), and we now cite and discuss these dynamics (lines 351-354).

5) For the observed decreased in midbrain neurons: is this observed across the whole midbrain, or is there selective decrease/increase in neurons selective to the social object (ie: those cells identified in Figure 1I)? What

are the anatomical connections here that link increased forebrain activity to suppressed midbrain/hindbrain? The authors claim there are directed projections from the forebrain to the midbrain and cite ref. 38 (Kunst et al), but there is no evidence for this. That paper states “For example, we found that the forebrain together with the preoptic region and the posterior tuberculum comprises a distinct module anatomically separated from the mid- and hindbrain, which seem. What are the authors referring to? Do they have evidence that forebrain excitation precedes midbrain inhibition on a single-trial basis?

We thank the Reviewer for these important points. In the revised manuscript, we now show that the decreased activity patterns in the midbrain and hindbrain are specific to the social cue (Figure 5A–C), and do not encode conspecific position (Figure S5C). While these midbrain and hindbrain neurons are spatially dispersed (Figure S4I), we have localized the telencephalic neurons with the opposite activity pattern within the pallium (Figure S6A). After confirming that our functionally identified neurons are pallial, we re-examined the telencephalic projections to the midbrain and hindbrain in the MapZebbrain atlas and found that they correspond to subpallial populations. We have therefore removed this point from the Discussion. Interestingly, although no established connectivity between these regions is known, targeted laser ablation of the small pallial cluster substantially reduced social approach relative to each fish's pre-ablation baseline (Figure 5F,G). Moreover, functional imaging one day post-ablation revealed loss of the brain-wide signature preceding approach movements (Figure 5H,I). Together, these findings strengthen our mechanistic understanding of the neural processes underlying social approach.

Minor points

1) The conspecific angular position on the observer's tectal map is a nice validation of social vision (Fig 1I), but the authors also say they see direction selectivity (line 117). Please show this map too - does this have any spatial order?

We have included a new figure presenting the direction selectivity map (Figure S1F). In agreement with previous reports of tectal direction-selective neurons (Hunter et al., 2013), this map shows no discernible spatial organization. The text was revised accordingly (lines 115-117).

2) For the argument about movement sequences of duration 9.5s: Given that the average movement rate is 5 movements per minute (Fig S1B), wouldn't that typically just be one movement? Most of the examples in Fig 2H are just showing 1 movement per “sequence”. How is this considered a sequence?

The sequence we refer to is the movement sequence of the conspecific. We now illustrate two example conspecific sequences to the left and right (Figure S3Hi), and provide the histogram of the duration from the onset of conspecific sequences to the onset of focal fish movements (Figure S4G).

3) the Y-axes in many plots are not matched to allow for fair comparisons (for instance, Fig S4C, Fig S4G, Fig 3B, Fig 3C,D,F,

Y-axes were matched to allow for fair comparisons.

4) when neurons are averaged or used for population-level analyses (decoding, cosine dist.), the authors chose a variety of different neurons (100 in some cases, 300 in others, 20 in other conditions). Why was this done? Please justify these choices in the methods section, or be consistent throughout.

We thank the Reviewer for this helpful comment. The differences in the number of neurons used across analyses reflect the varying population sizes identified in different brain regions. While hundreds of midbrain and hindbrain neurons showed reduced activity before approach movements (Figure 4A), only tens of telencephalic neurons exhibited the opposite pattern (Figure 4D). Accordingly, we selected the number of neurons used for decoding based on the size of these functionally defined populations (300 midbrain and hindbrain neurons, 30 telencephalic neurons). This is now explicitly stated in the Methods (lines 600-605). For visualization, we plotted the activity traces of the 100 most distinguishing midbrain and hindbrain neurons.

5) I was very surprised to see no description or discussion of neural activity in the dorsal thalamus – the main brain region identified for social attraction in zebrafish (see Kappel et al., Nature 2022, citation 28 here). As this is the other major study of neural imaging in response to social motion in zebrafish larvae/juveniles, I found this to be

a notable omission. If the authors did not see biological motion-sensitive neurons in the midline dorsal thalamus (which is within their imaging fields of view), then this major discrepancy deserves a substantial discussion.

The dorsal diencephalon was included in our imaging volume (Figure S1B, Methods). In this region, we observed frequency-tuned neurons that responded selectively to the characteristic bout frequency of conspecifics (Figure S6D), consistent with previous reports (Kappel et al., 2022). However, these neurons did not distinguish approach vs. non-approach movements. Signals preceding a movement that carried information about its nature arose in pallial, midbrain, and hindbrain populations. We now report these new diencephalic results and explicitly discuss alignment and limits of these results in the Discussion (lines 369-378).

6) given the maturation of the telencephalon between 6 and 14 days of age, how can the authors justify the brain atlas segmentation in Figure S5I? Anatomical structures present in their fish would not even be present in the 6dpf atlas. They may benefit from manual segmentation or alignment with an adult or juvenile brain atlas (see work from Elena Dreosti, Emre Yaksi, and others). At the very least, the authors should show examples of (or better yet, quantify) their atlas registration.

We would like to clarify that all 16 dpf brains were manually segmented, with a subset (n=6) registered to a 16 dpf reference fish from our dataset. An external atlas was used only to localize the identified telencephalic neurons within the pallium, and we now rely on the 21 dpf atlas (by Dreosti, as suggested by the Reviewer) as a developmental match. We have demonstrated this finding across fish (Figure S6A) and updated the Methods section accordingly (lines 447-457).

7) There is no evidence for the neural accumulation idea in the discussion (starting at line 338). I suggest they drop this conjecture, or back it up with evidence that punctate actions from the conspecific drives directed ramping or step-like accumulation to a boundary (movement initiation) in neural populations or single neurons.

As suggested by the Reviewer, we have removed this conjecture from the Discussion.

8) I am unsure about where the N values are coming from in some cases for statistics. In Fig 3B for example (but also 4B,E and S5), are these cells in an example fish, or pooled across all recorded fish? If the latter, how are the authors handling this nested data organization (non-independent cells within independent subjects) for statistical tests? For decoding accuracy and cosine distances, this is fine because there is one value per independent fish subject. Please clarify either way.

In general, neurons from different fish were not pooled for any statistical test, and the only exception was for atlas projections. Figure 3B shows neurons from a single fish. In some panels, we illustrate single-fish effects (e.g., Figures 4A and 4D), whereas for fish population-level analyses (e.g., Figures 4B and 4E) we used a single measure per fish, with the mean and SEM shown across all fish. We have added these clarifications to the figure captions where previously missing.

Overall, Lifshitz et al. tackle an interesting topic with new methods, but in many cases are missing details and justification for their conclusions. I think this manuscript could be a good fit for the wide readership of Nature Communications, but it would require substantial revisions and incorporation of the points raised above to be appropriate for publication.

We thank the Reviewer for the constructive comments. These additions strengthen validation of the novel assay, clarify data interpretation and methodology, and reinforce our conclusions.

Revision Outline

Manuscript NCOMMS-25-49635B, Lifshitz et al.

Overview of changes

We thank the Reviewers for their careful and constructive evaluation of our manuscript. We appreciate the positive feedback on the revised version. We have addressed all reviewer comments, performed additional control experiments, and revised the manuscript, figures, and supplementary material to improve experimental specificity, interpretational clarity, and conceptual framing. A brief summary of the main changes is provided below.

- **Added sham controls to validate pallial ablation.** We added an additional sham ablation experiment ($n = 10$), in which neurons in the olfactory bulb, an area not required for visually guided social behavior in this assay, were ablated instead of pallial neurons. This manipulation did not reduce social preference.
- **Clarified ablation details and specificity.** We now explicitly describe the ablation procedure in the Results section. In addition, we demonstrate that the ablation is spatially confined and selective, with neighboring neurons remaining functionally active and visual function preserved.
- **Refined conceptual framing.** We clarified our contribution to the spatiotemporal aspects of social behavior, acknowledged relevant prior work, and more carefully positioned our experimental social assay relative to VR-based approaches.
- **Corrected and tightened imaging-scope terminology.** We removed the term "brain-wide" throughout the manuscript and replaced it with precise descriptions of large-scale volumetric imaging, explicitly restricting conclusions to the imaged telencephalon, midbrain, and hindbrain.
- **Explicitly added caveats and contextual discussion.** We now clearly state limitations regarding pallial neuron identity and connectivity, as well as the reduced movement rate in the head-fixed condition.

In summary, we have addressed all reviewer concerns through additional experiments, analyses, and careful re-framing of methodological and conceptual claims. The revised manuscript now presents a technically transparent and conservatively framed account of distributed neural dynamics underlying social approach behavior, supported by appropriate controls and clearly articulated limitations.

Reviewer 1

The authors have addressed and incorporated all of the questions and suggestions in the revised version of the manuscript. At this point, I have nothing further to add or request. The results and the discussion has improved substantially, as specific brain nuclei were contextualized to explain the behavioural changes and the associated neural activity. The additional behavioural experiments (summarized in Figure 5) add value to the work and further support to the conclusions. I am very pleased with the final result.

We thank the Reviewer for their positive assessment and for the thoughtful comments throughout the review process, which substantially strengthened the manuscript.

Reviewer 2

My concerns were addressed by the authors with new experiments and analyses. I now recommend acceptance of the manuscript. However, I have two additional minor points that I think should be addressed before formally accepting the manuscript.

We thank the Reviewer for the positive assessment of the revised manuscript and for recommending its acceptance. We are pleased that the new experiments and analyses have addressed the Reviewer's concerns. Below, we address the two remaining minor points and describe the corresponding changes made in the manuscript.

1) The authors' characterization of the pallial neurons involved in social interactions has some limitations. Only their rough location in the pallial region was identified, and no information about their molecular identity, neurotransmitter content, and connectivity was provided. Moreover, random neurons in the pallium, and not specifically the ones active during social interactions, were targeted in cell ablation experiments. These limitations should be pointed out in the discussion.

We thank the Reviewer for this comment. We now explicitly acknowledge these limitations in the Discussion, including the coarse anatomical localization of pallial neurons and the lack of information regarding their molecular identity, neurotransmitter content, and connectivity (lines 369-373).

2) I would suggest to add information on number of samples used for the experiments in the figure legends, in addition to or instead of the main text.

We have now added the number of fish included in each analysis to the figure legends in addition to the main text.

Reviewer 3

The authors have submitted a substantially improved revision. The added behavioral analyses and clarifications are valuable. The clearer description of conspecific sequences and the more systematic examination of approach versus non-approach movements make the behavioral results more interpretable. The manuscript is clearer and the overall logic of the study is easier to follow. At the same time, there are significant concerns regarding (i) experimental controls for the ablation experiments, (ii) the internal consistency and interpretation of the neural analyses across brain regions, and (iii) the framing of the methodological and conceptual contributions. I outline these points below.

We thank the Reviewer for the careful evaluation of the revised manuscript and for the constructive assessment of the added behavioral analyses and clarifications. We are pleased that the Reviewer found the manuscript clearer and the behavioral results more interpretable. Below, we address each of the remaining concerns in detail, including those related to the ablation controls, the interpretation of neural analyses across brain regions, and the framing of the methodological and conceptual contributions.

1. Need for standard controls and experimental detail in the pallial ablation experiment

The inclusion of the pallial ablation experiment is potentially very informative and nicely complements the imaging data. However, in its current form, it does not yet meet standard expectations for zebrafish neural ablation studies.

First, the study lacks control (sham) ablations (e.g., applying the same laser protocol in a nearby region or in cells not implicated in approach behavior). For example, the habenula is not implicated in the present study and can potentially serve as a control region. Without such controls, it is not possible to determine whether the behavioral changes arise specifically from the loss of pallial neurons involved in social approach, as opposed to general consequences of the ablation procedure (local tissue damage, inflammation, stress, etc.). Second, the experimental details of the ablation should be more easily accessible directly from the Results. For the reader to evaluate the manipulation, it would be helpful to know approximately how many neurons are ablated per fish and how much residual activity remains in the targeted region (i.e. whether ablation leads to partial versus near-complete silencing). Third, because telencephalic structures lie close to the highly pigmented eyes, which efficiently absorb two-photon excitation light, there is a realistic risk that laser ablations could inadvertently affect the retina or optic pathways. A simple control for visual function would substantially increase confidence in the conclusions; without it, it remains difficult to exclude the possibility that some of the observed behavioral changes reflect unintended visual or other nonspecific effects.

We thank the Reviewer for these important comments. In response, we have added additional control experiments and clarified key experimental details of the pallial ablation in the Results. Specifically, we now include (i) a sham ablation control in a neighboring, non-socially relevant brain region in our visually guided social assay, (ii) quantitative characterization of the extent and spatial specificity of the pallial ablation, and (iii) explicit behavioral and neural controls demonstrating intact visual function following ablation. Together, these additions directly address concerns regarding procedural specificity, off-target effects, and unintended sensory impairments.

(i) Sham control for the ablation procedure

To control for behavioral changes that could arise from the ablation procedure itself, we included an additional sham control experiment, as suggested by the Reviewer. In this experiment, we applied the two-photon ablation protocol but targeted neurons in the olfactory bulb rather than the pallium, scattered in a comparable tissue volume. On average, a similar number of neurons (30 per fish) were ablated in the sham condition as in the pallial ablation. The olfactory bulb was selected as a control region due to its proximity to the pallium and because it is not required for visually driven social preference or approach behavior in this assay. We did not select the habenula as a control region because prior work indicates that habenular circuits encode stimulus valence and support goal-directed navigation across behavioral contexts (Palieri et al., *Current Biology*, 2024). Moreover, the habenula has been implicated in social behavior (Chou et al., *Science*, 2016), making it an unsuitable negative control. Fish subjected to olfactory bulb ablation ($n = 10$) did not show a reduction in social preference, indicating that the behavioral effect observed following pallial ablation is specific to the loss of pallial neurons rather than a nonspecific consequence of the ablation procedure. We now describe these results in Figure 5G, as well as in the Results (lines 317-320) and Methods (lines 635-641) sections.

(ii) Extent and specificity of the pallial ablation

To clarify the magnitude, spatial extent, and selectivity of the pallial ablation, we now explicitly describe the ablation parameters in the Results section (lines 311-313). On average, 46 pallial neurons were ablated per fish across both hemispheres, spanning a depth of 40 μm . Furthermore, we now show that this manipulation was spatially confined (Figure 5Ei), and functional imaging performed in the same pallial region one day post-ablation revealed that neighboring neurons remained functionally active (Figure 5Eii). Accordingly, the ablation did not silence the pallium globally but selectively targeted a sparse subset of pallial neurons.

(iii) Control for unintended visual or nonspecific effects

To address the concern that ablations near the eyes could inadvertently affect visual function, we now clarify that all ablated fish were assessed for successful prey capture to exclude gross visual or motor impairments, and only fish that successfully fed were included in subsequent experiments (lines 632-634). In addition, we now show that tectal neurons in pallium-ablated fish showed preserved selectivity to the conspecific's position, comparable to non-ablated fish (Figure S6C, lines 316-317, 634-635). Together, these behavioral and neural control analyses are consistent with the interpretation that the observed social deficits arise specifically from pallial neuron loss.

2. Consistency of analyses and interpretation across brain regions

The description of how different brain regions contribute to prediction of approach behavior remains somewhat dif-

difficult to reconcile across sections. In the section “Distinct neural processes predict upcoming approach movements on a single-trial level,” the emphasis is on midbrain and hindbrain as the regions from which upcoming movement identity (approach versus non-approach) can be robustly decoded from pre-movement population activity. In the subsequent section “Coordinated brain-wide neural dynamics precede approach behavior,” the narrative shifts to highlight the telencephalon, where neurons are described as showing the expected pattern of elevated activity before approach and reduced activity before non-approach. If telencephalic neurons show a pronounced and directionally consistent modulation, one might intuitively expect decoding performance in telencephalon not to be substantially lower than in midbrain or anterior hindbrain. However, comparing Fig. 3 and Fig. 4, telencephalon decoding accuracy appears substantially lower than in midbrain and hindbrain. The reason for this discrepancy is unclear from the current text and figures.

We thank the Reviewer for the opportunity to clarify this point. While decoding accuracy for approach and non-approach movements in the telencephalon (Figure 4F) appears lower than in the midbrain and hindbrain (Figure 3G), this difference is not statistically significant (Bonferroni-corrected t-test; telencephalon vs. midbrain: $p = 0.28$; telencephalon vs. hindbrain: $p = 0.24$). We have now included this statistical test in the legend of Figure 4G. Importantly, decoding accuracy reflects more than the mean magnitude or direction of neural modulation (Figure 4E). It is also influenced by trial-to-trial variability and the number of neurons available to the classifier. In our dataset, substantially fewer neurons were available for decoding in the telencephalon compared to the midbrain and hindbrain (30 vs. 300, lines 595-605), which likely constrains decoding performance despite clear and consistent population-level modulation. Thus, strong and directionally consistent telencephalic modulation can coexist with lower apparent decoding performance, without contradicting the interpretation that the telencephalon participates in coordinated pre-approach dynamics.

3. Use of “spatial factors”, “spatiotemporal integration”, and claims of novelty

The manuscript’s use of terms such as “spatial factors” and “spatiotemporal integration” remains somewhat imprecise, and some of the broader claims about novelty could be more carefully framed. For instance, the sentence, “The approach toward conspecifics is known to be influenced by various spatial factors, including the number of conspecifics, the visual load they impose on the retina, their shape and their perceived depth”, groups several quantities under “spatial factors.” While perceived depth is clearly spatial, “shape” is more naturally described as a visual stimulus feature, and other relevant dimensions such as color would also fall under “visual” rather than “spatial.” As written, the terminology blurs sensory categories in a way that may confuse readers about what exactly is meant by “spatial.” Moreover, several prior studies have already examined coordinated trajectories over time. When studies analyze the joint trajectories of two animals, they are by definition considering relationships that unfold across both space and time. In this context, suggesting that previous studies fail to jointly consider spatial and temporal aspects of social behavior is unnecessarily dismissive. The primary behavioral finding of the present study is that after a conspecific changes direction, the direction of the focal fish’s elicited movement depends on the egocentric position of the conspecific. This is an incremental but valuable refinement of current models of zebrafish social interaction. It would be more effective to state this contribution clearly, without relying on broad or potentially overreaching statements about prior literature.

We thank the Reviewer for this helpful clarification. We agree that our original wording grouped distinct sensory dimensions too broadly and could obscure the specific variables under study. We have therefore revised the text (lines 43-44) to restrict the term “spatial” to variables that explicitly describe the conspecific’s position in space (e.g., relative position and depth), and removed references to visual stimulus features (e.g., shape and retinal occupancy). We also agree that prior studies have examined coordinated trajectories over time, and we did not intend to imply otherwise. Accordingly, we now emphasize that our contribution does not lie in the general joint consideration of spatial and temporal factors of social behavior, but rather in specifically demonstrating how approach movements are jointly shaped by the relative position of a conspecific, the timing of its movements, and longer-timescale changes in its movement direction (lines 46-49).

On methodology: 4. Scope of imaging and use of “brain-wide” terminology

The authors now clearly specify that they image a $705 \times 352 \times 70 \mu\text{m}$ volume located approximately $110\text{--}180 \mu\text{m}$ below the skin. Given that the brain of a 16 dpf zebrafish is around $200 \mu\text{m}$ in depth, this implies that only a fraction of the total brain volume is sampled axially, and that the caudal part of the brain, including much of the hindbrain, lies outside the field of view seen in Supplementary Fig. 1. In this light, the repeated use of the term “brain-wide” is

problematic. Consider the two citations that the authors give in the rebuttal: “Across fish, 10,000-15,000 neurons were identified within this volume per fish, in line with previous brain-wide volumetric studies (Lin et al., 2020; Kappel et al., 2022).” First, in “Visual Recognition of Social Signals by a Tectothalamic Neural Circuit.” Nature 2022, Kappel et al. use a similar 2P functional imaging approach to the present study. Crucially, Kappel et al. never claim “brain-wide imaging” in the title, abstract, or results. Instead, they carefully describe the exact regions imaged and restrict their conclusions to the network that is actually sampled, explicitly avoiding overextending their methodological claims. By contrast, Lin et al. 2020 is a true brain-wide imaging study using light-field microscopy, with an imaging volume of $700\ \mu\text{m} \times 700\ \mu\text{m} \times 200\ \mu\text{m}$. This represents an order-of-magnitude more comprehensive sampling of the larval zebrafish brain than in the present work and approaches the common-sense meaning of “brain-wide” by attempting to cover most of the brain, rather than a relatively limited subvolume. Placing these two papers together under the umbrella of “previous brain-wide volumetric studies” conflates fundamentally different methodological regimes and obscures the care with which each prior study calibrated its claims to its actual coverage. The term “brain-wide” imaging is typically reserved for attempts to sample the brain at least somewhat comprehensively (for example, $700\ \mu\text{m} \times 700\ \mu\text{m} \times 200\ \mu\text{m}$), not for volumes that encompass only a small fraction of it. If such comprehensive coverage were attempted here, the authors could make a much stronger argument that the telencephalon signal they observe that ramp prior to approach movements is genuinely unique in the brain. However, if only roughly 10% of the brain is sampled, as in the present study, it remains entirely plausible that similarly tuned populations reside in unobserved regions, such as the caudal hindbrain or other omitted areas (e.g., hypothalamus). The real strength of the present study is that it records activity from several brain regions (telencephalon, midbrain, and an anterior portion of the hindbrain) at cellular resolution during a social assay. That is already substantial. To make the methodological claims more accurate and transparent, I strongly recommend removing the term “brain-wide” from the title and from strong statements in the abstract and discussion, unless the imaged volume is substantially expanded to cover the majority of the brain, and replacing “brain-wide” with more precise descriptions such as “multi-region volumetric imaging”. These changes would not diminish the value of the data; instead, they would increase the credibility of the methodological description.

We thank the Reviewer for this comment and for the careful discussion of imaging scope and terminology. As suggested, we have removed the term “brain-wide” from the title and replaced it throughout the manuscript with more precise descriptions of the imaged volume and the specific regions sampled (telencephalon, midbrain, and hindbrain). We now explicitly emphasize that our conclusions are restricted to these recorded regions and do not imply comprehensive sampling of the entire brain. We have also revised the abstract and discussion to avoid overextending methodological claims and to better highlight the central strength of the study: large-scale volumetric imaging at cellular resolution during a social behavior assay. We believe that this revised terminology and framing more accurately reflects the scope of the data and enhances the methodological transparency and credibility of the study.

5. Real conspecific assay versus VR / replayed video stimuli

The authors argue that their assay isolates the social visual cue and preserves the natural 3D kinematics and pose dynamics of a real conspecific, and that this “full visual complexity” is difficult to reproduce in VR. However, VR systems can present not only simplified synthetic stimuli such as dots, but also highly realistic stimuli, including videos or rendered 3D models of real fish. Therefore, contrasting the present assay with VR in terms of “full visual complexity” versus “simplified synthetic stimuli” is misleading, especially when referring to studies that have already used more sophisticated virtual conspecifics. Moreover, the current data do not isolate which specific aspects of “full visual complexity” (e.g. 3D kinematics, fine-scale posture, textural details) are needed for the observed differences in neural or behavioral responses. To present a more balanced and convincing methodological perspective, I suggest explicitly acknowledging that VR-based approaches can replay and systematically manipulate complex, realistic stimuli, including those derived from real conspecific movements, and reframing the comparison in terms of technical ease of implementation: the present assay provides a technically easy method to deliver naturalistic real-conspecific stimuli, while VR enables precise control and causal testing of stimulus components.

We agree that VR-based approaches can replay and systematically manipulate complex, highly realistic social stimuli, including videos or 3D renderings derived from real conspecific movements, and we did not intend to suggest otherwise. We have therefore revised the manuscript to avoid framing the comparison in terms of “full visual complexity” versus “simplified synthetic stimuli”. Instead, we now emphasize that VR-based assays and real-conspecific paradigms offer complementary strengths. VR enables the presentation of realistic social stimuli with precise experimental control, facilitating causal dissection of individual stimulus features. Our assay provides

a technically straightforward means of presenting naturalistic, real-conspecific visual input, without the need for stimulus rendering or calibration, under conditions compatible with stable, long-duration imaging (lines 336-338).

Reviewer 4

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

We thank Reviewer 4 for co-reviewing this manuscript and for their contribution to the evaluation process.

Reviewer 5

In this revised manuscript, the authors have improved upon their original manuscript, with new experiments, analyses, methodological clarifications, and discussion points. I believe the authors have adequately address the points I raised, and have correctly added context or caveats to the text when appropriate.

We thank the Reviewer for their thoughtful reassessment of the revised manuscript and for acknowledging the substantial improvements. We are pleased that the additions and clarifications adequately addressed the points raised. Below, we address the remaining minor points and describe the corresponding changes made in the manuscript.

In response to the Major points I raised: 1) The authors show, in FigS3G, that social approach is correlated between embedded and freely-moving conditions. While this not the strongest correlation, it is nonetheless significant for the low N, and I find this satisfactory. Though, I would prefer the authors not use the word “novel” to describe their assay in either the figures or the main text. Simply call it an “embedded”, or “tethered” assay (for instance, in the x-axis label of Fig S3G, and throughout the text and methods).

We thank the Reviewer for this constructive suggestion. We have removed the term “novel” throughout the manuscript and now refer to the assay as a tethered social assay, particularly when contrasting it with freely behaving conditions. Specifically, we updated the x-axis label in Figure S3G and revised the terminology in the main text and Methods accordingly.

2) This is a reasonable response – thank you for clarifying.

We thank the Reviewer for their positive assessment.

3) This explanation makes sense, but I think it is still a potential issue. Can the authors more explicitly state this caveat (much lower tethered movement rate) in the main text of manuscript, in addition to the methods (516-519)?

We thank the Reviewer for raising this point. We now explicitly state this caveat in the Results section, noting that the movement rate of the head-fixed focal fish is substantially lower than that of the freely moving conspecific (lines 96-97). This difference was previously described in the Methods and is now clearly highlighted in the main text as requested.

4) This is much clearer, especially showing the data in Figs 3F, S4D and S5E/F.

We thank the Reviewer for their positive assessment and are glad that the additional figures clarify this point.

5) I agree that the loss of mid/hind-brain activity after pallium laser ablation suggests there is a functional link between the forebrain and other brain regions, but there is no clear anatomical connectivity. Therefore, I appreciate the authors’ willingness to remove this speculation from their discussion.

We thank the Reviewer for this comment and for acknowledging our revision.

The authors also gave acceptable answers to the minor points, and I appreciated the nuance of the added discussion about thalamic populations, and discrepancy to the Kappel et al. study.

We thank the Reviewer for their positive assessment.

In sum, Lifshitz et al. have improved their manuscript in the revision – addressing my points and seemingly those of the other reviewers. If they can address any remaining points (including my minor additions above), I think this paper would be well suited for publications in Nature Communications.

We thank the Reviewer for their careful evaluation and positive assessment. We have addressed all remaining points raised in the review, including the minor additions noted above.

Revision Outline

Final revisions for Nature Communications

Manuscript NCOMMS-25-49635B, Lifshitz et al.

Overview of changes

We thank the Editor and the Reviewers for their thoughtful and constructive evaluation of our manuscript. We are grateful for the positive assessment of the revised version and, in this final revision, we have addressed the three remaining minor points raised by Reviewer 3. Specifically, we clarified sample-size differences across analyses, revised the reporting of extremely small p-values, and moderated the Discussion of the pallial ablation results to better reflect the scope of our conclusions. Below, we provide a point-by-point response.

Reviewer 1

We thank Reviewer 1 who was pleased with the previous version of the manuscript.

Reviewer 2

The authors address all my concerns in the revised manuscript. I recommend accepting it for publication in Nature Communications.

We thank the Reviewer for their careful evaluation of our manuscript and for their positive assessment of the revised version. We are pleased that the revisions have addressed all of their concerns.

Reviewer 3

The authors have satisfactorily addressed the major comments. The inclusion of sham controls is a particularly valuable and necessary addition, substantially strengthening the credibility of the pallial ablation results.

We thank the reviewer for their positive evaluation of the revised manuscript. We are pleased that the inclusion of sham controls addressed the major concerns and strengthened the rigor and credibility of the pallial ablation results.

Minor comments

1. Variations in sample size across figures and analyses are occasionally puzzling. For example: Why does 3G include 18 fish, whereas 3H includes 20? Please clarify the exclusion criteria used for each analysis more explicitly in the results section/figure legends. In the behavioral analyses, the sample size is sometimes reported as $n = 44$, whereas in other cases six animals are excluded, resulting in $n = 38$. While the Methods indicate that these exclusions correspond to animals that did not exhibit measurable social preference, it would improve clarity to explicitly remind the reader of this rationale at the relevant points in the Results, or to clearly indicate which analyses include or exclude these animals. It would also be helpful to briefly note whether inclusion of these non-social animals qualitatively affects the reported behavioral effects.

To improve clarity, we now explicitly report the sample size for each analysis in the figure legends, as requested by the Editor. All exclusion criteria are described in the Methods section (lines 435-438, 445-448, 534-535, 551-552, 607-608). Importantly, as mentioned in the Methods (lines 436-438), these exclusions do not affect the reported behavioral effects or their statistical significance.

2. For p-values that are effectively zero, reporting exact numerical values (e.g. $p = 6 \times 10^{-174}$ in Fig. S3) is not meaningful.

Extremely small p-values are now reported as $p < 0.001$ throughout the manuscript and figure legends.

3. Finally, I have some reservations about the framing of the pallial ablation results in the Discussion. The manuscript states that the loss of predictive dynamics suggests a role for pallial activity in the distributed neural signature preceding social approach behavior. While this interpretation is plausible, it may be somewhat overstated. The authors didn't specifically target neurons with approach activity. Thus while the ablating semi-random cells within the telencephalon clearly disrupts telencephalic function, the resulting behavioral and neural effects are not necessarily specific to social approach per se. Alternative explanations, such as impairments in working memory, spatial representation, or other cognitive processes that indirectly affect social behavior, cannot be excluded. I would suggest rewording this section to acknowledge these possibilities and to more explicitly frame the conclusions as reflecting a requirement for intact pallial processing rather than a selective role in social approach alone.

To avoid overstating the specificity of the pallial ablation results, we have softened the wording in the Discussion and now frame this conclusion as reflecting a requirement for intact pallial processing for the distributed neural signature preceding social approach behavior (lines 372-374). We also note in the manuscript that the ablated neurons were not specifically selected based on their approach-related activity, and that their functional role will require further investigation (lines 370-372).

Reviewer 4

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

We thank Reviewer 4 for co-reviewing this manuscript and for their contribution to the evaluation process.

Reviewer 5

In this revised manuscript, the authors have addressed the remaining issues I raised in the last round, and found their responses to other reviewers' comments reasonable. I have no further comments.

We thank the Reviewer for their careful evaluation of the revised manuscript. We are pleased that the remaining issues have been satisfactorily addressed, and we appreciate their thoughtful assessment of both our revisions and our responses to the other reviewers' comments.