Email

Manuscript #	eLife-RP-RA-2024-103097
Title	Canonical neurodevelopmental trajectories of structural and functional manifolds
Corresponding Author	Alicja Monaghan (University of Cambridge)
Date:	28th Dec 24 05:06:35
Last Sent:	28th Dec 24 05:06:35
Created By:	Redacted
From:	editorial@elifesciences.org
To:	Alicja.Monaghan@mrc-cbu.cam.ac.uk
BCC:	Redacted
Subject:	Reviews for your submission to eLife: eLife-RP-RA-2024-103097
Email	Dear Dr Monaghan, Thank you for submitting your article "Canonical neurodevelopmental trajectories of structural and functional manifolds" for consideration by eLife. Your article has been reviewed by 2 peer reviewers, and the assessment has been overseen by a Reviewing Editor and Andre Marquand as the Senior Editor. First, please accept my apologies for the earlier mix-up we had with this manuscript and for the delay in returning the decision to you. As you may know, I had difficulties finding suitable reviewing editors to review your work, but I have now obtained sufficient input to prepare an eLife assessment (see below). In short, while the reviewers saw value and novelty in the results presented, they feel that additional sensitivity analyses are necessary to validate the methodological decisions made and thereby better support the claims in the manuscript. Full details can be found in the attached reviews. --- eLife Assessment This important study provides insights into the neurodevelopmental trajectories of structural and functional connectivity gradients in the human brain and their potential associations with behaviour and psychopathology. While certain aspects of the methodology are rigorous, the evidence supporting the findings is currently incomplete and would benefit from additional sensitivity analyses to evaluate methodological choices supporting the findings. This study will be of interest to neuroscientists interested in understanding functional connectivity across development. --- Public Reviews: Reviewer #1 (Public review): Summary:

In this study, the authors advance our understanding of neurodevelopmental changes in the brain's structural and functional connectivity, as well as their coupling. The paper presents evidence of alterations in and stability of the principal organizational gradients of structure and function across development (age) and contrasts them between neurotypical and neurodivergent individuals. The authors further extend their findings by exploring links with graph theory measures of brain connectivity and indices of nodal structure-function coupling. Finally, the developmental shifts in structural and functional brain organization are examined for potential associations with cognitive and psychopathological markers. The results suggest that structure-function coupling, both brain-wide and within specific functional networks, is associated with certain cognitive dimensions but not with measures of psychopathology.

Strengths:

This manuscript makes a significant contribution to the field by synthesizing previous research while offering novel insights into the developmental trajectories of brain organization. A key strength of this study lies in its integration of both structural and functional connectivity data, providing a comprehensive view of brain changes throughout development. The authors present findings that challenge earlier reports of shifts in principal gradients during late childhood and early adolescence (e.g., Dong et al., 2021; Xia et al., 2022), underscoring an important inconsistency that could have broader implications for our understanding of developmental brain reorganization. The introduction and discussion sections are well-crafted, offering a thorough review of relevant prior studies and effectively situating the current findings within the broader context of the literature. Additionally, the study design and methodology are detailed and adhere to recommended best practices, demonstrating a commendable level of rigor in the formulation of the study and its various assessments.

Weaknesses:

Despite these strengths, I think there are aspects of the manuscript that would benefit from further refinement. Below is detailed feedback and suggestions provided point-by-point.

- Lack of Sensitivity Analyses for some Key Methodological Decisions:
Certain methodological choices in this manuscript diverge from approaches used in previous works. In these cases, I recommend the following: (i) The authors could provide a clear and detailed justification for these deviations from established methods, and (ii) supplementary sensitivity analyses could be included to ensure the robustness of the findings, demonstrating that the results are not driven primarily by these methodological changes. Below, I outline the main areas where such evaluations are needed:
> Use of Communicability Matrices for Structural Connectivity Gradients: The authors chose to construct structural connectivity gradients using communicability matrices, arguing that diffusion map embedding "requires a smooth, fully connected matrix." However, by definition, the creation of the affinity matrix already involves smoothing and ensures full connectedness. I recommend that the authors include an analysis of what happens when the communicability matrix step is omitted. This sensitivity test is crucial, as it would help determine whether the main findings hold under a simpler construction of the affinity matrix. If the results significantly change, it could indicate that the observations are sensitive to this design choice, thereby raising concerns about the robustness of the conclusions. Additionally, if the concern is related to the

large range of weights in the raw structural connectivity (SC) matrix, a more conventional approach is to apply a log-transformation to the SC weights (e.g., log(1+$SCij$)), which may yield a more reliable affinity matrix without the need for communicability measures.
> Individual-Level Gradients vs. Group-Level Gradients: Unlike previous studies that examined alterations in principal gradients (e.g., Xia et al., 2022; Dong et al., 2021), this manuscript focuses on gradients derived directly from individual-level data. In contrast, earlier works have typically computed gradients based on grouped data, such as using a moving window of individuals based on age (Xia et al.) or evaluating two distinct age groups (Dong et al.). I believe it is essential to assess the sensitivity of the findings to this methodological choice. Such an evaluation could clarify whether the observed discrepancies with previous reports are due to true biological differences or simply a result of different analytical strategies.
> Procrustes Transformation: It is unclear why the authors opted to include a Procrustes transformation in this analysis, especially given that previous related studies (e.g., Dong et al.) did not apply this step. I believe it is crucial to evaluate whether this methodological choice influences the results, particularly in the context of developmental changes in organizational gradients. Specifically, the Procrustes transformation may maximize alignment to the group-level gradients, potentially masking individual-level differences. This could result in a reordering of the gradients (e.g., swapping the first and second gradients), which might obscure true developmental alterations. It would be informative to include an analysis showing the impact of performing vs. omitting the Procrustes transformation, as this could help clarify whether the observed effects are robust or an artifact of the alignment procedure. (Please also refer to my comment on adding a subplot to Figure 1)
> SC-FC Coupling Metric: The approach used to quantify nodal SC-FC coupling in this study appears to deviate from previously established methods in the field. The manuscript describes coupling as the "Spearman-rank correlation between Euclidean distances between each node and all others within structural and functional manifolds," but this description is unclear and lacks sufficient detail. Furthermore, this differs from what is typically referred to as SC-FC coupling in the literature. For instance, the cited study by Park et al. (2022) utilizes a multiple linear regression framework, where communicability, Euclidean distance, and shortest path length are independent variables predicting functional connectivity (FC), with the adjusted R-squared score serving as the coupling index for each node. On the other hand, the Baum et al. (2020) study, also cited, uses Spearman correlation, but between raw structural connectivity (SC) and FC values. If the authors opt to introduce a novel coupling metric, it is essential to demonstrate its similarity to these previous indices. I recommend providing an analysis (supplementary) showing the correlation between their chosen metric and those used in previous studies (e.g., the adjusted R-squared scores from Park et al. or the SC-FC correlation from Baum et al.). Furthermore, if the metrics are not similar and results are sensitive to this alternative metric, it raises concerns about the robustness of the findings. A sensitivity analysis would therefore be helpful (in case the novel coupling metric is not similar to previous ones) to determine whether the reported effects hold true across different coupling indices.

- Methodological ambiguity/lack of clarity in the description of certain evaluation steps: Some aspects of the manuscript's methodological descriptions are ambiguous, making it challenging for future readers to fully reproduce the analyses based on the information provided. I believe the following sections would benefit from additional detail and clarification:
> Computation of Manifold Eccentricity: The description of how eccentricity was computed (both in the results and methods sections) is unclear and may be

problematic. The main ambiguity lies in how the group manifold origin was defined or computed. Specifically:
(1) In the results section, it appears that separate manifold origins were calculated for the NKI and CALM groups, suggesting a dataset-specific approach.
(2) Conversely, the methods section implies that a single manifold origin was obtained by somehow combining the group origins across the three datasets, which seems contradictory.
Moreover, including neurodivergent individuals in defining the central group manifold origin is conceptually problematic. Given that neurodivergent participants might exhibit atypical brain organization (as suggested by Fig. 1), this inclusion could skew the definition of what should represent a typical or normative brain manifold. A more appropriate approach might involve constructing the group manifold origin using only the neurotypical participants from both the NKI and CALM datasets. Given the reported similarity between group-level manifolds of neurotypical individuals in CALM and NKI, it would be reasonable to expect that this combined origin should be close to the origin computed within neurotypical samples of either NKI or CALM. As a sanity check, I recommend reporting the distance of the combined neurotypical manifold origin to the centers of the neurotypical manifolds in each dataset. Moreover, if the manifold origin was constructed while utilizing all samples (including neurodivergent samples) I think this needs to be reconsidered.
> Computation of SC-FC coupling: As noted in a previous comment, the explanation of this procedure is vague. The description lacks detail on the specific steps taken and differs from previous standard approaches in the field. I suggest clarifying the methodology and comparing with previous SC-FC coupling metrics.
> Performing Procrustes transformation: The brief explanation in the first paragraph of page 30 does not provide enough information about the procedure or its justification. Since the Procrustes transformation alters the shape of individual gradients, it could artificially inflate consistency across development. I recommend including a rationale for using the Procrustes transformation and conducting a sensitivity analysis to assess its impact on the findings. Additionally, clarifying how exactly the transformation was applied to align gradients across hemispheres, individuals, and or datasets would help resolve ambiguity.

- Insufficient Supporting Evaluations for Certain Claims:
There are instances where additional analyses are necessary to substantiate the claims made in the manuscript. Without these evaluations, some conclusions may be premature or potentially misleading. I believe the following points need further analysis or, alternatively, adjustments to the claims:
> Evaluating the Consistency of Gradients Across Development: The results shown in Fig. 1.e are used as evidence suggesting that gradients are consistent across ages. However, I believe additional analyses are required to identify potential sources of the observed inconsistency compared to previous works. The claim that the principal gradient explains a similar degree of variance across ages does not necessarily imply that the spatial structure of the gradient remains stable. The observed variance explanation is hence not enough to ascertain inconsistency with findings from Dong et al., as the spatial configuration of gradients may still change over time. Moreover, the introduction of the Procrustes transformation (not used by Dong et al.) further ambiguates the cause of this inconsistency. I suggest the following additional analyses to strengthen this claim: (1) Alignment to Group-Level Gradients: Assess how much of the variance in individual FC matrices is explained by each of the group-level gradients (G1, G2, and G3, for both FC and SC). This analysis could be visualized similarly to Fig. 1.e, with age on the x-axis and variance explained on the y-axis. If the explained variance varies as a function of age, it may indicate that the gradients are not as

consistent as currently suggested. (2) For each individual's gradients (G1, G2, and G3, separately for FC and SC, without Procrustes transformation), evaluate their spatial similarity to the corresponding group-level gradients using a similarity metric (e.g., correlation coefficient). High spatial similarity, without a Procrustes transformation, would support the claim of stable gradient structures across development. On the other hand, if the similarities alter during development (e.g. such that at a certain age, individual G1 is less similar to group G1) this would contradict the stability of gradients during development. These additional analyses could potentially be included as additional panels in Fig. 1. In case significant deviations are observed, it might help refine the interpretation of the results and provide a more nuanced understanding of developmental changes in gradient organization.
> Prediction vs. Association Analysis: The term "prediction" is used throughout the manuscript to describe what appear to be in-sample association tests. This terminology may be misleading, as prediction generally implies an out-of-sample evaluation where models trained on a subset of data are tested on a separate, unseen dataset. If the goal of the analyses is to assess associations rather than make true predictions, I recommend refraining from using the term "prediction" and instead clarifying the nature of the analysis. Alternatively, if prediction is indeed the intended aim (which would be more compelling), I suggest conducting the evaluations using a k-fold cross-validation framework. This would involve training the Generalized Additive Mixed Models (GAMMs) on a portion of the data and testing their predictive accuracy on a held-out sample (i.e., different individuals). Additionally, the current design appears to focus on predicting SC-FC coupling using cognitive or pathological dimensions. This is contrary to the more conventional approach of predicting behavioral or pathological outcomes from brain markers like coupling. Could the authors clarify why this reverse direction of analysis was chosen? Understanding this choice is crucial, as it impacts the interpretation and potential implications of the findings.

- Methodological considerations
> In typical applications of diffusion map embedding, sparsification (e.g., retaining only the top 10% of the strongest connections) is often employed at the vertex-level resolution to ensure computational feasibility. However, since the present study performs the embedding at the level of 200 brain regions (a considerably coarser resolution), this step may not be necessary or justifiable. Specifically, for FC, it might be more appropriate to retain all positive connections rather than applying sparsification, which could inadvertently eliminate valuable information about lower-strength connections. Whereas for SC, as the values are strictly non-negative, retaining all connections should be feasible and would provide a more complete representation of the structural connectivity patterns. Given this, it would be helpful if the authors could clarify why they chose to include sparsification despite the coarser regional resolution, and whether they considered this alternative approach (using all available positive connections for FC and all non-zero values for SC). It would be interesting if the authors could provide their thoughts on whether the decision to run evaluations at the resolution of brain regions could itself impact the functional and structural manifolds, their alteration with age, and or their stability (in contrast to Dong et al. which tested alterations in high-resolution gradients).

- The Issue of Abstraction and Benefits of the Gradient-Based View:
> The manuscript interprets the eccentricity findings as reflecting changes along the segregation-integration spectrum. Given this, it is unclear why a more straightforward analysis using established graph-theory measures of segregation-integration was not pursued instead. Mapping gradients and computing eccentricity adds layers of

abstraction and complexity. If similar interpretations can be derived directly from simpler graph metrics, what additional insights does the gradient-based framework offer? While the manuscript argues that this approach provides "a more unifying account of cortical reorganization," it is not evident why this abstraction is necessary or advantageous over traditional graph metrics. Clarifying these benefits would strengthen the rationale for using this method.

Reviewer #2 (Public review):

Summary:

This study aims to show how structural and functional brain organization develops during childhood and adolescence using two large neuroimaging datasets. It addresses whether core principles of brain organization are stable across development, how they change over time, and how these changes relate to cognition and psychopathology. The study finds that brain organization is established early and remains stable but undergoes gradual refinement, particularly in higher-order networks. Structural-functional coupling is linked to better working memory but shows no clear relationship with psychopathology.

Strengths:

This study effectively integrates two different modalities (structural and functional) to identify shared patterns. It is supported by a relatively large dataset, which enhances its value and robustness.

Weaknesses:

General Comments:
- The introduction is overly long and includes numerous examples that can distract readers unfamiliar with the topic from the main research questions.

- While the methods are thorough, it is not always clear whether the optimal approaches were chosen for each step, considering the available data.
Detailed Comments:
- The use of COMBAT may have excluded extreme participants from both datasets, which could explain the lack of correlations found with psychopathology.
- Some differences in developmental trajectories between CALM and NKI (e.g., Figure 4d) are not explained. Are these differences expected, or do they suggest underlying factors that require further investigation?
- There is no discussion of whether the stable patterns of brain organization could result from preprocessing choices or summarizing data to the mean. This should be addressed to rule out methodological artifacts.

Recommendations for the authors: please note that you control which revisions to undertake from the public reviews and recommendations for the authors.

Reviewer #1 (Recommendations for the authors):

- Other minor thoughts and comments
> The manuscript should clarify from the outset that the reported sample size (N) includes multiple longitudinal observations from the same individuals, and does not reflect the number of unique participants.
> The term "structural gradients" is ambiguous in the introduction. Clarify that these gradients were computed from structural and functional connectivity matrices, not from other structural features (e.g., cortical thickness).
> The use of "streamline counts" is misleading, as the method uses SIFT2-weighted fiber bundle capacity rather than raw streamline counts. It would be better to refer to this measure as "SIFT2-weighted fiber bundle capacity" or "FBC".
> Page 5: The sentence, "We calculated the normalized angle of each structural and functional connectome to derive symmetric affinity matrices," is unclear and needs clarification.
> Figure 1.c: A nonparametric permutation test (e.g. Mann-Whitney U test) could quantitatively identify regions with significant group differences in nodal gradient values, providing additional support for the qualitative findings.
> Figure 1.a: "Affine A" likely refers to the affinity matrix. The term "affine" may be confusing; consider using a clearer label. It would also help to add descriptive labels for rows and columns (e.g., region x region).
> Figure 1.d: Are the cross-group differences statistically significant? If so, please indicate this in the figure.
> Figure 2.c: Consider adding plots showing changes in eccentricity against (1) degree centrality, and (2) weighted local clustering coefficient. Additionally, a plot showing the relationship between age and mean eccentricity (averaged across nodes) at the individual level would be informative.
> Figure 2.b, Considering the results of the following sections, it would be interesting to include additional KDE/violin plots to show group differences in the distribution of eccentricity within different 7 functional networks.
> Figure 3: Several panels lack axis labels for x and y axes. Adding these would improve clarity.
> Figure 3: It would be helpful to include a plot showing the GAMM predictions versus real observations of eccentricity (x-axis: predictions, y-axis: actual values).
> Figure 4.b: Why were ranks shown instead of actual coefficient of variation values? Consider including a cortical map visualization of the coefficients in the supplementary material.
> It is unclear whether the statistical tests finding significant dataset effects are capturing effects of neurotypical vs. neurodivergent, or simply different scanners/sites. Could the neurotypical portion of CALM data also be added to distinguish between these two sources of variability affecting dataset effects (i.e. ideally separating this to the effect of site vs. neurotypicality would better distinguish the effect of neurodivergence).
> The statement that "differences between datasets only emerged when taking development into account" seems inaccurate. Differences in eccentricity are evident across datasets even before accounting for development (see Fig 2.b and the significance in the Scheirer-Ray-Hare test).
> The handling of longitudinal data by adding a random effect for individuals is not clear in the main text. Mentioning this earlier could be helpful.
> The 30mm threshold for filtering short streamlines in tractography is uncommon. What is the rationale for using such a large threshold, given the potential exclusion of many short-range association fibers?
> Given the spatial smoothing of fMRI data (6mm FWHM), it would be beneficial to apply connectome spatial smoothing to structural connectivity measures for consistent

spatial smoothness.
> The sentence "whose connectomes were successfully thresholded" in the methods is unclear. What does "successfully thresholded" mean? Additionally, this seems to be the first mention of the Schaefer 100 and Brainnetome atlas; clarify where these parcellations are used.
> Why was harmonization performed only within the CALM dataset and not across both CALM and NKI datasets? What was the rationale for this decision?
> In the harmonization methodology, it is mentioned that "If harmonisation was successful, we'd expect any significant effects of scanner type before harmonisation to be non-significant after harmonisation." However, given that there were no significant effects before harmonization, the results reported do not help in evaluating the quality of harmonization.
> The exclusion of subcortical areas from connectivity analyses is not justified.
> In the kNN imputation method, were uniform weights used, or was an inverse distance weighting applied?

Reviewer #2 (Recommendations for the authors):

-- The results section is extensive, with a large number of reports, while the discussion is relatively short and lacks in-depth analysis of the findings. Moving some results into the discussion could help balance the sections and provide a deeper interpretation.
- Although imputing missing data was necessary, it would be useful to compare results without imputed data to assess the impact of imputation on findings.
- As mentioned before, some differences in developmental trajectories between CALM and NKI (e.g., Figure 4d) are not explained. Are these differences expected, or do they suggest underlying factors that require further investigation?

PUBLISHING THE REVIEWED PREPRINT:

1) The Reviewed Preprint (including the full text of the preprint we reviewed, the eLife assessment, and public reviews) will typically be published in two weeks' time.

* Please let us know as soon as possible if you would like to provide provisional author responses to accompany the first version of the Reviewed Preprint: these should be sent by email within the next two weeks; they should not be uploaded to the system. If you will need slightly longer to provide provisional author responses, please let us know, so we can schedule things accordingly.

* Provisional author responses need only address the public reviews, for example by outlining what revisions are planned. However, full author responses, when you come to resubmit, should address both sets of comments (public reviews and recommendations for the authors).

* If there are any factual errors in the eLife assessment or public reviews, or other issues we should be aware of, please let us know as soon as possible.

2) After publication of the Reviewed Preprint, you can use the link below to submit a

revised version. There is no deadline to resubmit (but please do not resubmit within the next two/three weeks). Please note that we are always willing to consider the first round of revisions; requests to consider subsequent rounds of revision will need to be justified (in an attempt to avoid multiple rounds of review).

Before resubmitting, please ensure that you update the preprint at the preprint server to correspond with the revised version. When resubmitting, please also upload a response to the previous letter and tracked changes version of the article file, alongside the updated files. Upon submitting a revised version, we will ask the editors and reviewers if it's appropriate to update their assessment and public reviews, which will be included alongside the revised Reviewed Preprint. At that time we will also post the recommendations to the authors and the author responses you provide with the revised version. In the complete author response that accompanies the revised version, please respond to the public reviews and the recommendations to the authors.

3) Alternatively, you can proceed with the current version of the Reviewed Preprint (once published), without revisions, and request an eLife Version of Record. See the Author Guide for further information: https://elife-rp.msubmit.net/html/elife-rp_author_instructions.html#vor. However, the vast majority of authors request a Version of Record after a single round of revisions.

4) After publication of eLife's Reviewed Preprint, you also have the option to submit/publish in another journal instead: if you choose to do this, please let us know so we can update our records, but please note that the first version of the Reviewed Preprint will remain online.

Please note: Reviewed Preprints will be indexed by Google Scholar, OpenAlex and Scopus. The final version, the Version of Record (VOR), will be indexed by PubMed and other databases. A partial feed of VORs (representing around 93% of articles) is sent to Web of Science. Read more: https://elifesciences.org/inside-elife/c11c6101. Please contact us if you would like to discuss this further.

You can use the following link to access your manuscript, and then submit a revised version when you are ready (or proceed to a Version of Record, as appropriate):

[Link Not Available

Best wishes,

Andre Marquand
Reviewing Editor, eLife

Andre Marquand
Senior Editor, eLife

---

	at the address 95 Regent Street, Cambridge, CB2 1AW, UK. -----