BIAS-ACCOUNTING META-ANALYSES OVERCOME CEREBELLAR NEGLECT TO REFINE THE CEREBELLAR BEHAVIORAL TOPOGRAPHY

Neville Magielse[1-3*], Aikaterina Manoli[1,2,4,5], Simon B. Eickhoff[1,3], Peter T. Fox[6,7], Amin Saberi[1-3#] & Sofie L. Valk[1-3#,*]

[1] *Institute of Neuroscience and Medicine (INM-7: Brain and Behaviour), Research Center Jülich, Jülich, Germany.*

[2] *Otto Hahn Cognitive Neurogenetics Group, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany.*

[3] *Institute of Systems Neuroscience, Medical Faculty and University Hospital Düsseldorf, Heinrich Heine University, Düsseldorf, Germany.*

[4] *Minerva Fast Track Group Milestones of Early Cognitive Development, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany.*

[5] *Faculty of Medicine, Leipzig University, Leipzig, Germany.*

[6] *Research Imaging Institute, University of Texas Health Science Center at San Antonio, San Antonio, TX, USA.*

[7] *Biggs Institute for Alzheimer's and Neurodegenerative Diseases, University of Texas Health Science Center at San Antonio, San Antonio, TX, USA.*

[#] **AS** and **SLV** contributed to the manuscript equally and share last authorship.

Corresponding authors: Neville Magielse (n.magielse@fz-juelich.de) and Sofie L. Valk (s.valk@fz-juelich.de)
Permanent address: Max Planck Institute for Human Cognitive and Brain Sciences, Dr. Sofie L. Valk, Stephanstraße 1A, 04103 Leipzig, Saxony, Germany.
Telephone: +49 341 9940-2658 | Fax: +49 341 9940-104

NM. Conceptualized the manuscript; assisted in code writing and analysis; interpreted results; created figures; wrote the manuscript; incorporated manuscript reviews.

AM. Assisted in manuscript conceptualization; provided manuscript reviews (several occasions); assisted in data interpretation.

SBE. Assisted in manuscript conceptualization; provided technical support; provided revision comments manuscript.

PTF. Provided the BrainMap data; provided manuscript review.

AS*. Conceptualized the manuscript; wrote code; performed chief analysis; assisted in method writing; provided manuscript reviews (several occasions).

SLV*. Assisted in manuscript conceptualization; provided manuscript reviews (several occasions); assisted in data interpretation; provided supervision.

Abbreviations: *ALE = activation likelihood estimation; BD = behavioral domain; CBMA = coordinate-based meta-analysis; C-SALE = cerebellum-specific activation likelihood estimation; GM = gray matter; MACM = meta-analytic connectivity-modeling; MDTB = multi-domain task battery; SA = spatial autocorrelations; SCALE = specific coactivation likelihood estimation; SNR = signal-to-noise ratio; SUIT = spatially unbiased infratentorial.*

Abstract

The cerebellum plays important roles in motor, cognitive, and emotional behaviors. Previous cerebellar coordinate-based meta-analyses and mappings have attributed different behaviors to cerebellar subareas, but an accurate behavioral topography is lacking. Here, we show overrepresentation of superior activation foci, which may be exacerbated by historical cerebellar neglect. Unequal foci distributions render the null hypothesis of standard activation likelihood estimation unsuitable. Our new method, cerebellum-specific activation-likelihood estimation (C-SALE), finds behavioral convergence beyond baseline activation rates. It does this by testing experimental foci versus null models sampled from a data-driven, biased probability distribution of finding foci at any cerebellar location. Cerebellar mappings were made across five BrainMap task domains and thirty-five subdomains, illustrating improved specificity of the new method. Twelve of forty (sub)domains reached convergence in specific cerebellar subregions, supporting dual motor representations and placing cognition in posterior-lateral regions. Repeated subsampling revealed that whereas action, language and working memory were relatively stable, other behaviors produced unstable meta-analytic maps. Lastly, meta-analytic connectivity modeling in the same debiased framework was used to reveal coactivation networks of cerebellar behavioral clusters. In sum, we created a new method for cerebellar meta-analysis that accounts for data biases and can be flexibly adapted to any part of the brain. Our findings provide a refined understanding of cerebellar involvement in human behaviors, highlighting regions for future investigation in both basic and clinical applications.

Introduction

The cerebellum has received considerable attention due to its involvement in modulating activity of wide-spread regions of the brain[1–5]. This perspective is consistent with current neuroscientific paradigms that conceptualize the brain as a complex array of interconnected networks rather than discrete, isolated areas[6]. Although the cerebellum has traditionally been associated with motor control, consensus is now that it plays integrative roles in brain-wide networks spanning behavioral domains (BDs) including emotion, language, and social cognition across development and aging[7–13].

The cerebellar circuit is organized into relatively separated functional modules that are connected to diverse extracerebellar areas in reciprocal loops[14–23]. This evolutionarily highly conserved organization[24–31] might facilitate cerebellar involvement in various human behaviors and brain networks. This is speculated to involve very similar computations at the algorithmic level[32–38] (although see compelling arguments for cerebellar algorithmic diversity[39,40]). Cerebellar functional topography is hence suggested to largely mirror its connectivity[18,24,41]. The chief cerebellar computation likely involves calculating error signals by mismatching predicted behavioral outcomes with actual sensory input[42–44]. This error signal then alters extracerebellar activity, a process which over time facilitates learning of smooth movements and other behaviors[3,38]. The cerebellum has massive bandwidth, with fifty billion granule cells responding to approximately forty million neocortical neurons (a thousand-to-one ratio)[39,45,46]. It responds to changing inputs through modulation of neuronal excitability at a manifold of LTD- and LTP-modulated synapses[47], including prominently LTP at Purkinje cell-parallel fiber synapses[33–37,48]. This, alongside the repetitive nature[1] of its circuit and connections to the basal ganglia and cerebrum[17,49,50], make the cerebellum an exemplary structure for (supervised) learning of highly specific representations across behaviors[51,52].

Decisively mapping cerebellar functions to subregions is, however, challenging. First, the small cerebellum (10% of brain volume) and its many folds[25,45,53] exaggerate partial volume effects at relatively crude neuroimaging resolutions. Spatial inaccuracies are magnified in the case of cerebellar spatial misalignment[54], compounding interindividual variability in cerebellar behavioral mapping[55,56]. Secondly, the signal-to-noise ratio (SNR) in the cerebellum is generally low, with common physiological noise and artifacts due to the location in the head and proximity to blood vessels[57]. This makes it difficult to reliably obtain functional signals. Lastly, blood-oxygen level-dependent (BOLD) functional magnetic resonance imaging (fMRI)

in the cerebellum may mostly reflect inputs from the mossy fiber system[39,57–59], further complicating its meaning for cerebellar function relative to the rest of the brain.

Several strategies to overcome these issues exist. First, high-resolution mapping in few individuals can reveal accurate, individualized resting-state topographies[55]. Next, deep task data across multiple individuals can map behaviors to subregions, dividing the cerebellum into non-overlapping regions. These intend to be maximally homogenous internally, and maximally different to others. Twenty-six different task activations across twenty-four individuals revealed how cerebellar functions readily cross lobular boundaries[60]. This parcellation effectively maps cerebellar functions, as revealed by high intra-region homogeneity and inter-region difference[60,61]. Moreover, adding large-scale resting-state data can improve discriminability of regional borders[56]. Functional fusion[62] with task-based data improved inferences made solely from resting-state data[63,64]. Lastly, high-quality task-based mapping of specific behaviors replicated the established double motor representation and additionally revealed three separate working memory, social cognition, and language mappings[65]. Together, these (among other) parcellations and mappings have revealed much of the known cerebellar functional topography. They paint a somewhat coherent picture: motor representations are consistently located in both anterior and posterior-inferior regions, whereas cognitive functions tend to occupy posterior-lateral regions such as crura I-II, but also inferior lobules IX-X. However, many aspects differ substantially: across maps different behaviors may be included or omitted, and the same cerebellar subregion ascribed quite different behaviors.

Crucially, although parcellations reveal relevant aspects of cerebellar behavioral mapping, creating a comprehensive behavioral topography involves combining task data from much larger samples. Mapping rather than parcellation approaches are preferred as they allow behavioral overlaps and do not impose full cerebellar coverage. Notwithstanding, for synthesis of task-based functional imaging findings into a common behavioral map, one needs to overcome the challenges noted above. Coordinate-based meta-analysis (CBMA) approaches such as activation likelihood estimation (ALE)[66–69] fit this purpose. Specifically, ALE finds brain regions that are activated above chance, implying being consistently involved, in a task or task domain. In other words, when activity in a task (domain) converges despite imaging inaccuracies, this implies highly consistently activity in that region. Repeated meta-analyses across random subsets of experiments within a task (domain) can then be used to understand how consistently experiments support the mapping.

CBMAs, both at the whole-brain-level and using the cerebellum as region-of-interest (ROI), have provided complementary understanding of cerebellar functions. Studies at the whole-brain-level have provided insight on whether the cerebellum and its subregions are activated above chance in functional processes. In this way, cerebellar involvement in, for example, audition[70,71], verbal working memory[72], and cognitive control[73,74] has been established. CBMAs within the cerebellar ROI have enabled localization of behaviors to cerebellar subareas, including motor (learning) behavior, emotion, and several aspects of social and executive cognition[75–82]. CBMAs have been used to map cerebellar topography more comprehensively, providing indication that diverse motor and cognitive functions occupy distinct cerebellar subregions[76]. Comparable analyses have been performed within the socio-cognitive domain at larger scale[78,79]. Recognizing that cerebellar behavioral topography directly reflects brain-wide connectivity, meta-analytic connectivity modeling (MACM) revealed coactivation networks of two structural cerebellar subdivisions[83], inspired by the primate literature[84]. Moreover, clustering of MACM maps has been used to parcellate the cerebellum into behaviorally relevant communities[85].

Yet, despite this extensive work, several issues persist. First, small sample sizes and focus on specific behaviors or task domains limit interpretation from early meta-analyses. Next, structural cerebellar subdivisions were used as connectivity seeds in MACM[83,85]. However, lobules may have limited utility in defining functional borders, which readily cross lobular boundaries[60]. A recently published meta-analytic atlas[86] of NeuroSynth activation coordinates[87] addresses these issues, but three remain open. First, though NeuroSynth is the largest available dataset of brain-wide activation coordinates, labeling of experimental contrast was done automatically and no sample size information was available. Therefore, the smaller, but manually curated BrainMap database[88] is preferred. Secondly, any atlas assumes that every cerebellar voxel is involved in just one function. Contrasting, the mapping studies discussed above indicate that overlap of cerebellar involvement across behaviors is expected. Hence, accurate cerebellar mapping should not be limited by this constraint. Thirdly and crucially for the current study, the meta-analytic atlas does not address the marked cerebellar reporting biases in the neuroimaging literature[83,86,89] that need to be overcome for accurate behavioral mapping of the cerebellar cortex.

Specifically, a historical experimental design choice has seen the cerebellum often not being retained in the field-of-view (FOV) of functional scans. Shifting the FOV to frontal regions

improves data quality locally, at the cost of cerebellar coverage[77,79,83,89]. Likely connected to this historical neglect a substantial, accentuated bias has made its way into the neuroimaging literature[86]. Critically, foci are strongly skewed to superior cerebellar locations. Biased neglect appears to have exacerbated the inherently unequal distributions of foci across the brain. This means that the classic implementation of ALE may readily report convergence in overreported, superior cerebellar regions. Convergence in underreported, inferior regions is much less likely. Often, one of two strategies have been adapted in cerebellar CBMAs: **1.** include all data and report convergence for the biased sample of foci; or **2.** manually confirm full cerebellar coverage. Unfortunately, the former option overestimates convergence in oversampled regions (and vice versa). The latter, far more common strategy, leads to more accurate but possibly smaller-sampled or costly CBMAs, due to excessive manual work and the loss of many potentially eligible experiments. Additionally, it still does not guarantee that foci will be distributed equally.

In sum, despite the deep and diverse literature, an accurate and comprehensive meta-analytic map of cerebellar activations across BDs is lacking. Crucially, this map provides answers to the following questions: **(1)** *For a given behavior, what cerebellar subregion(s) show activity convergence above chance relative to the rest of the cerebellum?* and **(2)** *For a given behavior, what area(s) of the brain are coactivated with the cerebellar subregion(s) [found in **(1)**] above chance?* In this study, we create a novel implementation of ALE, here called cerebellum-specific ALE (C-SALE), that updates its null hypothesis to account for reporting biases. In this way, we can incorporate all data regardless of coverage, nonetheless overcoming literature reporting biases. The new method aims to overcome the limitations of both strategies discussed above. To create the map, we considered whole-brain studies, examining the largest set of cerebellar foci (many non-cerebellar-focused studies nonetheless report it), and recognizing cerebellar activations in relation to whole-brain activations. By then treating the cerebellum as ROI and constructing the null model within it, we can answer the main question of our study: where in the cerebellum is convergence of activity for each behavior elevated above baseline cerebellar activations?

Results

Study inclusion

We identified behavioral datasets by querying the BrainMap database for task-based fMRI and positron emission tomography (PET) data in healthy adults in February 2024, creating eligible

datasets for five BDs) and thirty-five subdomains. These domains refer to the division of tasks into human-understandable subcategories by the BrainMap team. Each subdomain includes tasks describing a distinct behavioral construct, and belongs to a more global BD (*Action, Cognition, Emotion, Interoception, Perception*)[88,90]. Note that a single task or experiment can map to several subdomains, and with that, BDs. 1,129 unique studies (16,410 participants; 4,783 cerebellar foci) were included. For the overall number of experiments, coordinates, and subjects included in each BD and subdomain C-SALE analysis, see **Supplementary File 1**.

Assessing improvements of C-SALE over classic ALE

Previous work has used ALE to map activations in behavioral (sub)domains to cerebellar subareas. The regular ALE method assumes equal spatial distributions of reported effects (foci). However, visualizing locations of cerebellar foci reported in the BrainMap dataset compared to voxels along the z-axis illustrates substantial spatial biases (**Figure 1a**). Foci were strongly skewed towards superior regions. Along the x- and y-axes, distributions of foci largely mirrored those of cerebellar voxels (**Supplementary Figure 1a, b**). It appeared that vermal foci were somewhat overrepresented (**Supplementary Figure 1a**). Whereas classic ALE assumes spatial homogeneity of foci, C-SALE considers the observed, unequal probabilities of finding foci at any voxel (**Figure 1b**). This baseline cerebellar probabilistic foci distribution, generated from domain-general foci reported in the BrainMap dataset, revealed two substantial probability hotspots in bilateral anterior lobules. Inferior-posterior regions had lowest probabilities, and the distribution was generally left-right symmetrical. Note that unequal distributions of reported effects are not unique to the cerebellum, but common across the brain (**Supplementary Figure 2)**.

Comparing ALE and C-SALE mappings

Whereas classic ALE constructs null models by sampling truly random cerebellar gray matter (GM) coordinates, C-SALE uses the probability distribution to sample pseudorandom GM coordinates, weighted by the probability distribution. In essence, comparing peak activation coordinates within a behavioral (sub)domain to this null model tests where in the cerebellum activity converges beyond baseline cerebellar activity. Comparing BD maps for classic ALE (**Figure 1c**) and C-SALE (**Figure 1d**) revealed improvements of our new method. First, whereas in ALE most of the superior half of the cerebellum reached convergence invariant of BD, in C-SALE convergence was only reached in specific locations for Action. Action reached convergence in bilateral lobules V-VI and right VIIIa, b. Examining unthresholded z-maps for

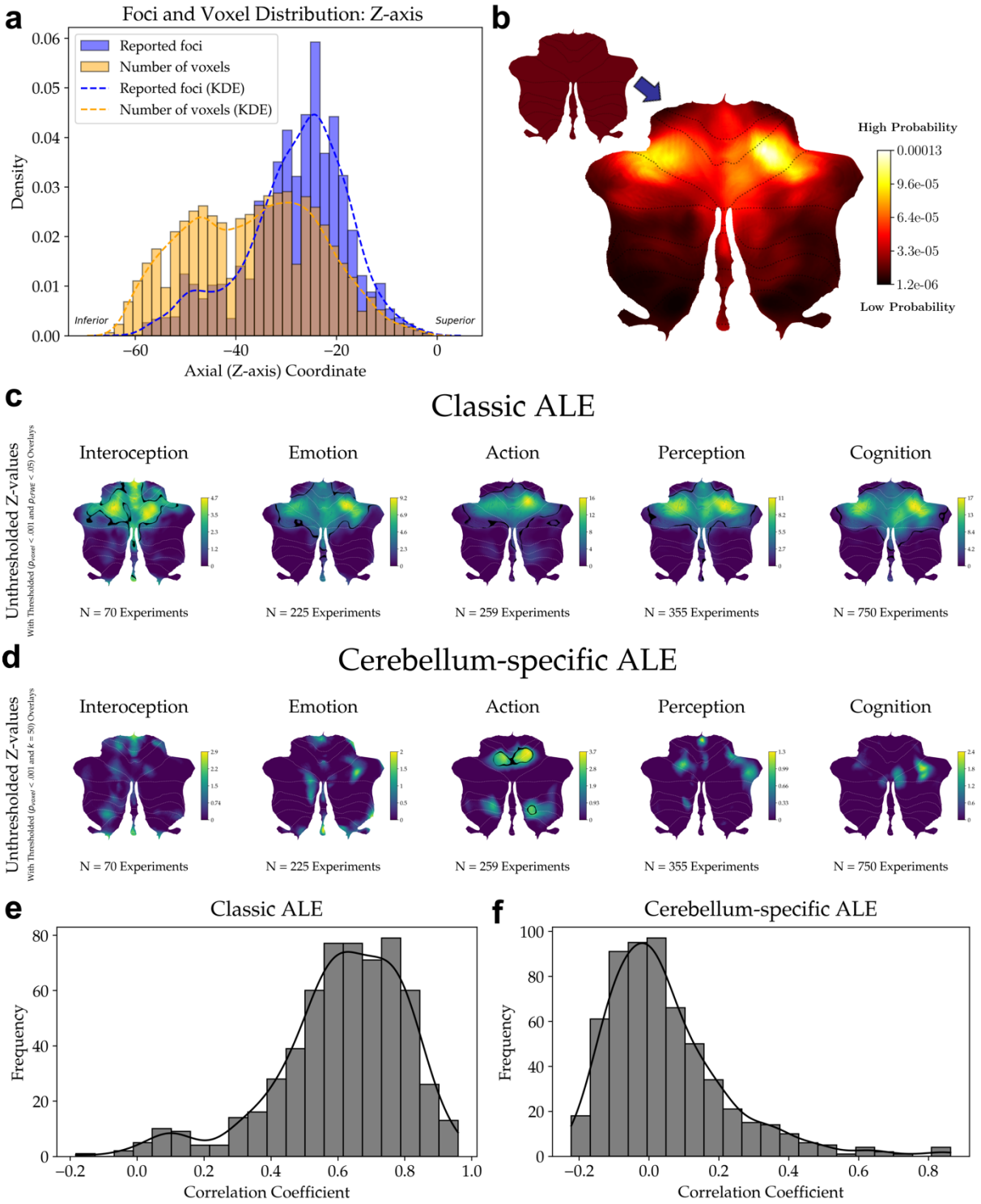


Figure 1: Improvements of the new cerebellum-specific ALE method over classic ALE for cerebellar meta-analysis. Methodological assessment of cerebellum-specific activation-likelihood estimation (C-SALE). (**a**) Maps the distribution of reported effects versus the number of cerebellar voxels across the Z-axis (superior (Z = 0) to inferior (Z = -70)). (**b**) Illustrates the null hypotheses of classic ALE (smaller flatmap) and C-SALE (larger flatmap). Whereas classic ALE assumes spatial homogeneity of reported effects (foci), C-SALE uses the heterogeneous distribution of foci found across all experiments in the BrainMap database as null assumption. Voxel- (C-SALE) and cluster-wise (classic ALE) statistical significance was determined by comparing observed values versus 10,000 random foci distributions sampled from either distribution. (**c, d**) Show differences between resulting meta-analytic maps from classic ALE (**c**) and C-SALE (**d**). Unthresholded z-maps are

shown with a black outline of the clusters that reached convergence (p_{voxel} < .001 and k = 50 for C-SALE and $p_{cFWE.}$ < .05 using a height threshold of p_{voxel} < .001 for classic ALE) overlaid. Behavioral domains (BDs) are ordered by sample size from small to large. Note that in (**d**), only specific cerebellar subregions, including an inferior region, reach convergence (within Action). (**e, f**) Show distributions of correlation coefficients between pairs of *subdomain x subdomain* unthresholded z-maps, for classic ALE (**e**) and C-SALE (**f**), respectively. Note that the median correlation for regular ALE is .63, whereas for C-SALE it is .01. Abbreviations: KDE = kernel density estimation.

BDs illustrate peaks of (subthreshold) convergence in inferior regions (**Figure 1c, d**). Second, we compared spatial correlations of unthresholded maps for pairs of BDs and subdomains. Whereas some spatial correlation should be expected across behaviors, excessive correlations indicate a lack of specificity. ALE subdomains illustrate such excessive (median = .63) correlations (**Figure 1e**). C-SALE normalizes the correlations (median = .01) (**Figure 1f**). Full hierarchically clustered correlation heatmaps for domains are reported in **Supplementary Figure 3a, b** and heatmaps for subdomains are reported in **Supplementary Figure 4.** Both illustrate that whereas classic ALE finds high correlations across combinations, C-SALE finds high correlations between related behaviors only. As we established the improved performance of C-SALE, we continue to report results for this method. In turn, the BD maps reported in **Figure 1d** also represent BD-level results.

Cerebellum-Specific Activation Likelihood Estimation (C-SALE)

Next, we use C-SALE to find if and where activations in each behavioral (sub)domain converged. C-SALE results for BDs (**Figure 1d**; **Figure 2a**) and subdomains (**Figure 2b, Figure 3**) are reported on the cerebellar flatmap. For both, unthresholded maps are plotted alongside an outline of thresholded (p_{voxel} < .001 and k = 50) clusters (**Figure 1d**; **Figure 3**). Binary locations of convergence for BDs are illustrated in **Figure 2a**, and **Figure 2b** illustrates how subdomain convergence maps to the BD they are organized under in BrainMap. For BDs, full unthresholded and thresholded (p_{voxel} < .001 and k = 50) z-maps are reported in **Supplementary Figure 4a-c**. Full z-maps for subdomains are provided in **Supplementary Figures 6-8.** For convenience, we report cluster sizes (in mm^3) and peak coordinates (in MNI ICBM-152 space[91]) for every C-SALE analysis in **Supplementary File 1**.

Behavioral subdomain topography of the cerebellum

Behavioral subdomains aim to capture more specific behavioral constructs. Hence, after investigating BDs, we investigated cerebellar subregional convergence in subdomains. Convergence was more readily reached in subdomains (eleven out of thirty-five; **Figure 3**). As

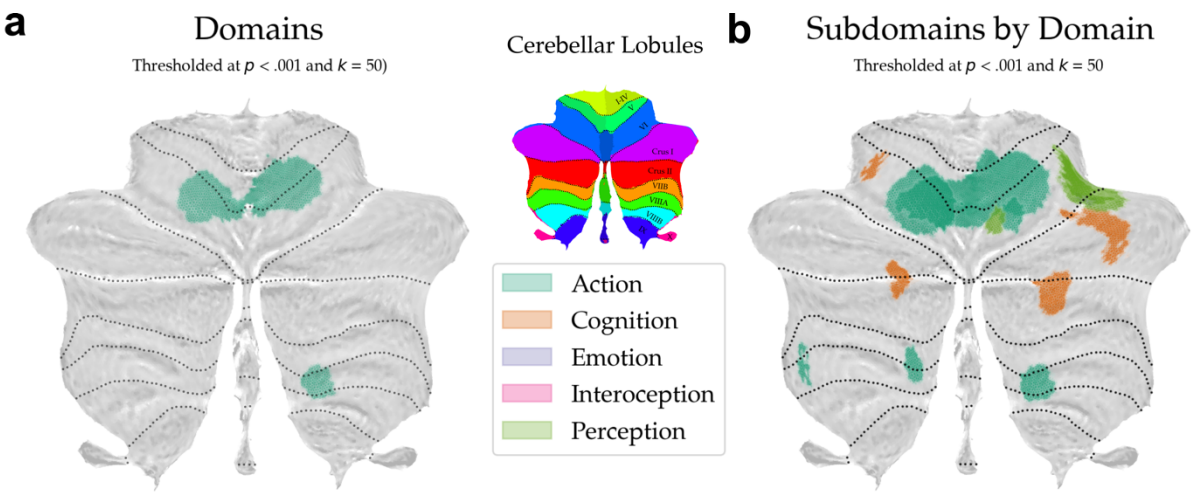


Figure 2: Locations of convergence across behavioral domains. (**a**) Summary of converging cerebellum-specific ALE maps for behavioral domains (BDs). Binary locations of convergence (p_{voxel} < .001 and k = 50) are plotted to a common flatmap. Only Action reached convergence. The next panel shows the cerebellar lobular definition of Larsell[54,92,93] (top) and a color mapping for BDs (bottom). Both legends apply to **a** and **b**. In (**b**), subdomain results (see **Figure 3**) are summarized by BD. Binary locations of converging subdomains, colored by the BD they belong to, are plotted onto a common flatmap to highlight spatial overlap.

Figure 2b illustrates, these subdomains belonged to Action, Cognition, Emotion, and Perception. Clusters were spread relatively equally across left-right and inferior-superior regions. Subdomains within a BD tended to map to distinct cerebellar subregions. Accordingly, subdomains were first ordered by the BD they belong to, then by sample size (**Figure 3**). Within Action, Observation, Speech Execution, and Execution reached convergence. Observation converged in two clusters in left VIIb and VIIIa, Speech Execution in left V-VI and right VI, and Execution in bilateral (including vermal) V-VI and right VIIb-VIIIb. Next, within Cognition, Social Cognition, Working Memory, and Language converged. Social Cognition converged in bilateral crura I-II, Working Memory in right VI-crus I, and Language in left VI. Sadness was the only Emotion subdomain to converge, in right VI-crus I. Within Perception, Vision - Motion converged in right VI, whereas Vision - Shape and Vision overall converged in right V-crus I. Lastly, Somesthesis converged in left VI.

Spatial stability of C-SALE maps

Then, to assess how consistently behaviors map to cerebellar subareas, we performed a complementary set of repeated subsampling analyses. Specifically, we reran C-SALE analyses in fifty random subsets of experiments, comparing spatial correlations between each pair of subsample C-SALE z-maps within a (sub)domain. To account for the variability in the number

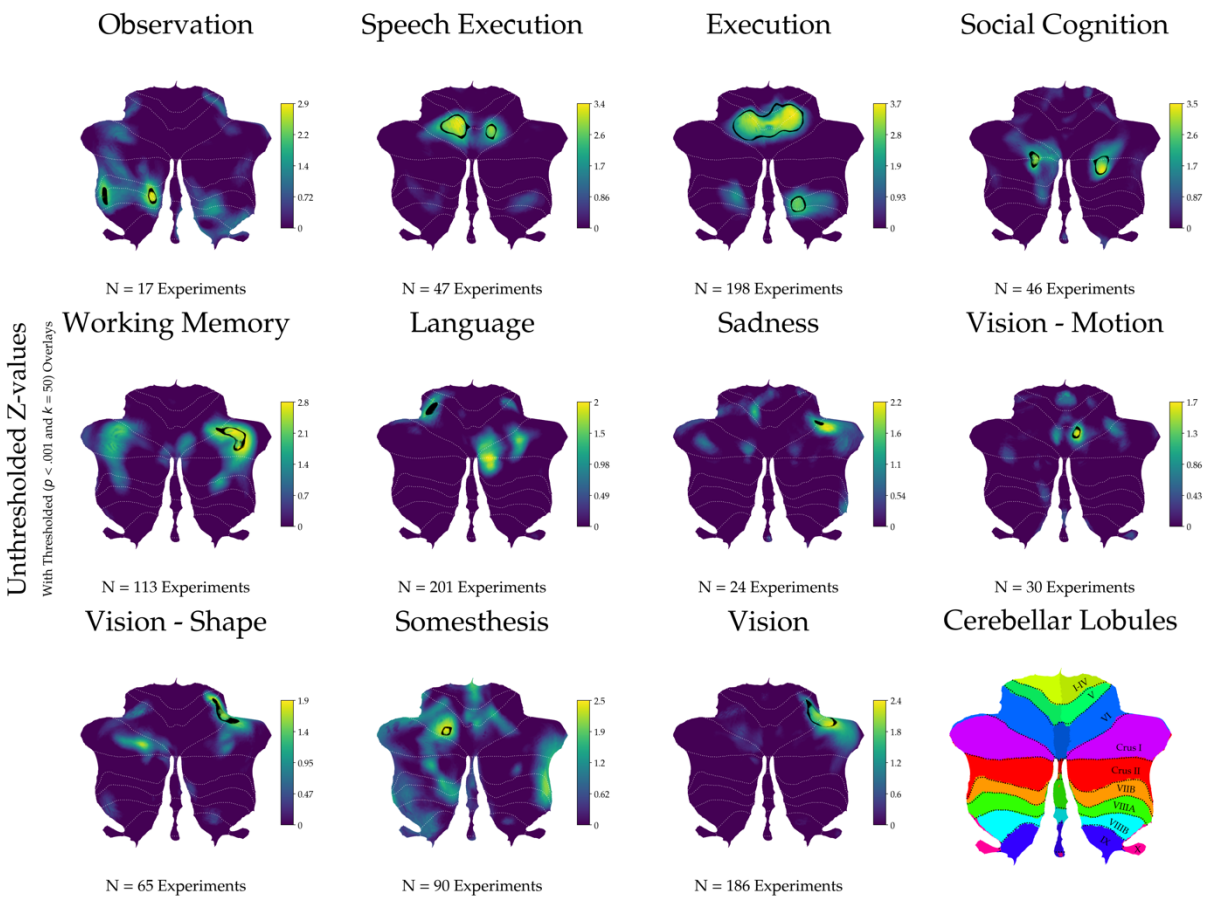


Figure 3: Cerebellum-specific ALE maps for behavioral subdomains. Unthresholded z-maps are shown with a black outline of the locations that reached convergence ($p_{voxel} < .001$ and k = 50) overlaid. Subdomains are first ordered by behavioral domain (alphabetically) and then by sample size (from small to large). The last panel shows the cerebellar lobular definition of Larsell, digitized into neuroimaging space by Diedrichsen and colleagues[54,92,93].

of experiments across (sub)domains, two complementary subsampling strategies were used for both BDs and subdomains: **1)** n = 50 per subsample; and **2)** $n_{subsample} = 0.2 * n_{(sub)domain}$. For BDs, we additionally illustrated how stability develops along a range of absolute or proportional subsample sizes. The repeated subsampling datasets correspond to those reported in **Supplementary File 1** but were only run for (sub)domains with $n_{experiments} \geq 100$, leaving five of five BDs and seventeen of thirty-five subdomains.

Spatial consistency of unthresholded maps

Spatial correlations of unthresholded subsample maps revealed that most BDs (**Figure 4a, b**) and subdomains (**Figure 4d, e**) were moderately stable. First, BD mapping stability naturally increased as sample sizes increased (**Figure 4a, b**). This was partially driven by increasing proportions of overlapping experiments across subsamples, illustrated by the large increase in mapping stability between n = 25 and n = 50 subsampling for Interoception relative to other

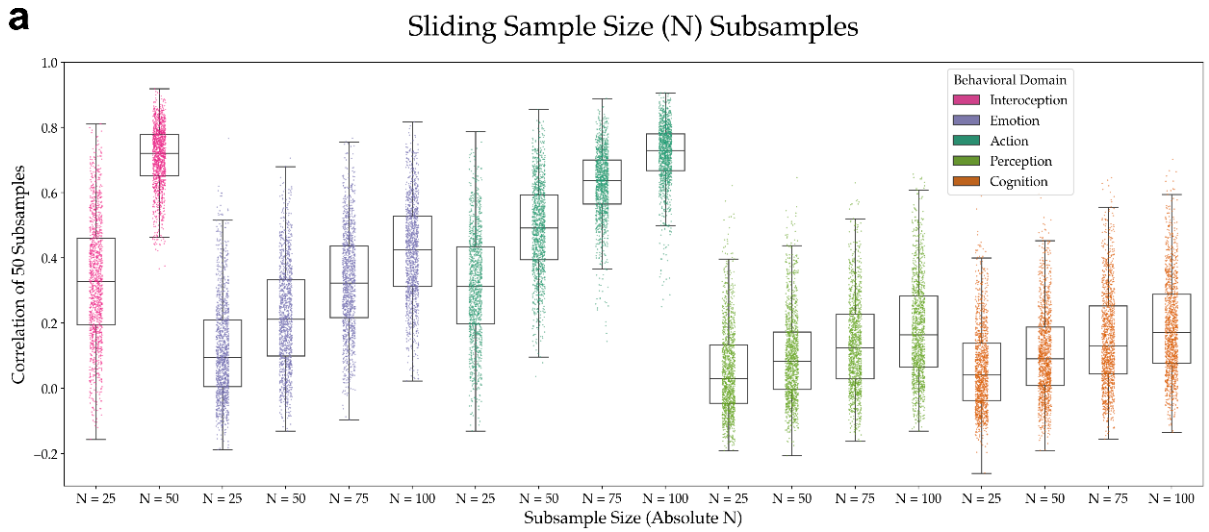
BDs (**Figure 4a**). Varying the subsampling proportion allowed for comparison of mapping stability between BDs. Stability increased as subsampling proportions increased, generally doing so similarly across BDs (**Figure 4b**). Cognition and especially Action stood out for high stability: at .8 subsampling proportions, their median spatial correlations were .87 (SD ± .06) and .94 (SD ± .03), respectively. For Interoception (.81; SD ± .07), Emotion (.78; SD ± .08), and Perception (.70; SD ± .11) median spatial correlations were somewhat lower. For the full list of median correlations and SDs per configuration, see **Supplementary Files 2-5**.

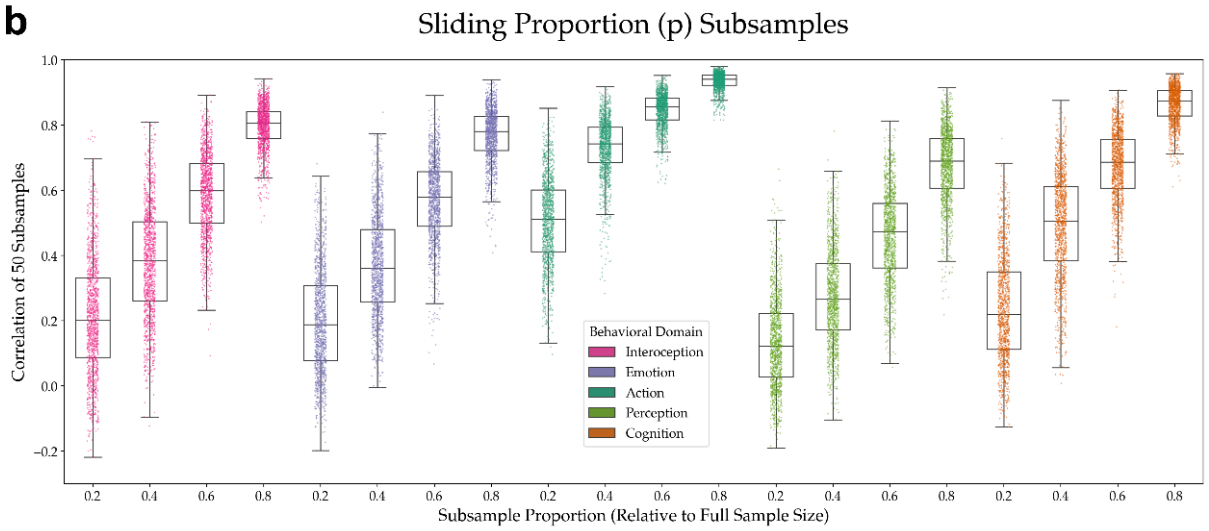
Consistency of voxel-wise cerebellar convergence

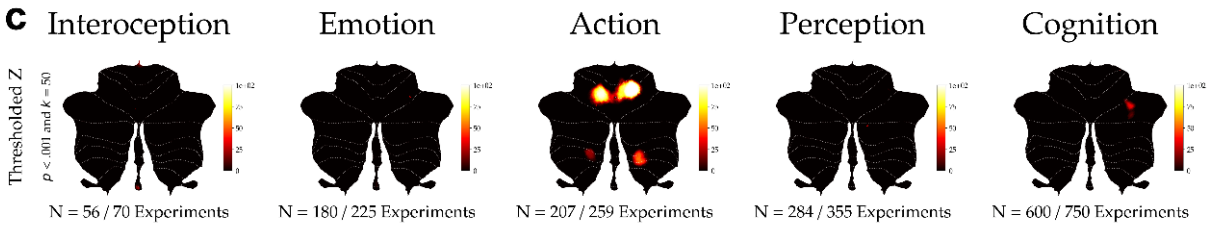
We also assessed how consistently convergence exceeded the threshold ($p_{voxel} < .001$ and k = 50). Specifically, we mapped subsamples per (sub)domain to a common flatmap, visualizing proportions of subsamples that reached convergence at each voxel for BDs (**Figure 4c, Supplementary Figure 9**) and subdomains (**Supplementary Figure 10**). Focusing on BDs, whereas unthresholded maps were rather comparable (based on relatively high spatial correlations (**Figure 4b**)), thresholds were only reached consistently in Action. Here, many voxels were significant across all subsamples for the .6 and .8 proportion subsampling (**Figure 4c, Supplementary Figure 9c)**. Even at the low .2 subsampling proportion, more than half of the subsamples reached convergence in many voxels (**Supplementary Figure 9a**). In other BDs, much fewer subsamples reached convergence, congruent with not reaching convergence in the main analyses (**Figure 4c, Supplementary Figure 9)**. Perception and Cognition reached convergence in a substantial percentage of subsamples for the .6 (peak voxel = 26% of subsamples for both) and .8 (32% and 36%, respectively) subsampling proportions. Notably, within Emotion and Interoception, fewer subsamples reached convergence at the .8 proportion relative to the .6 proportion, indicating that these maps were highly unstable. These tendencies were similarly illustrated by subdomains. For full voxel-wise peak percentages of significant subsamples across (sub)domains, as well as sampling sizes and proportions, see **Supplementary File 6**.

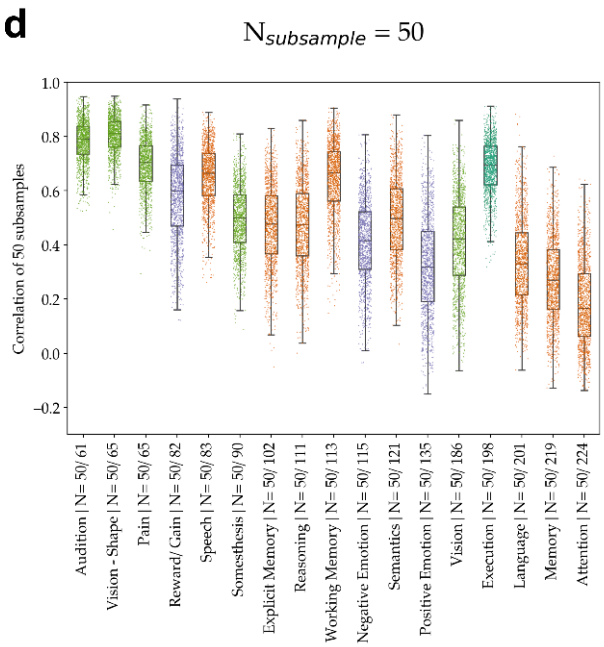
Correspondence of C-SALE maps to published cerebellar parcellations and mappings

To contextualize our meta-analytical cerebellar mappings, we compare them with several established mappings and parcellations, including the cerebellar multi-domain task-battery (MDTB)[60], mid-granularity hierarchical atlas[56], seven-network resting-state atlas[63], and functional gradients[94] (**Figure 5**). Note that for continuous comparisons, spatial correlations were calculated between every *C-SALE map* and *target map*. For dichotomous parcellations,


a
Sliding Sample Size (N) Subsamples
Correlation of 50 Subsamples
Behavioral Domain
Interoception
Emotion
Action
Perception
Cognition
N = 25
N = 50
N = 75
N = 100
Subsample Size (Absolute N)


b
Sliding Proportion (p) Subsamples
Correlation of 50 Subsamples
Behavioral Domain
Interoception
Emotion
Action
Perception
Cognition
Subsample Proportion (Relative to Full Sample Size)


c
Interoception
Emotion
Action
Perception
Cognition
Thresholded Z
p < .001 and k = 50
N = 56 / 70 Experiments
N = 180 / 225 Experiments
N = 207 / 259 Experiments
N = 284 / 355 Experiments
N = 600 / 750 Experiments


d
$N_{subsample} = 50$
Correlation of 50 subsamples
Audition | N= 50/ 61
Vision - Shape | N= 50/ 65
Pain | N= 50/ 65
Reward/ Gain | N= 50/ 82
Speech | N= 50/ 83
Somesthesis | N= 50/ 90
Explicit Memory | N= 50/ 102
Reasoning | N= 50/ 111
Working Memory | N= 50/ 113
Negative Emotion | N= 50/ 115
Semantics | N= 50/ 121
Positive Emotion | N= 50/ 135
Vision | N= 50/ 186
Execution | N= 50/ 198
Language | N= 50/ 201
Memory | N= 50/ 219
Attention | N= 50/ 224

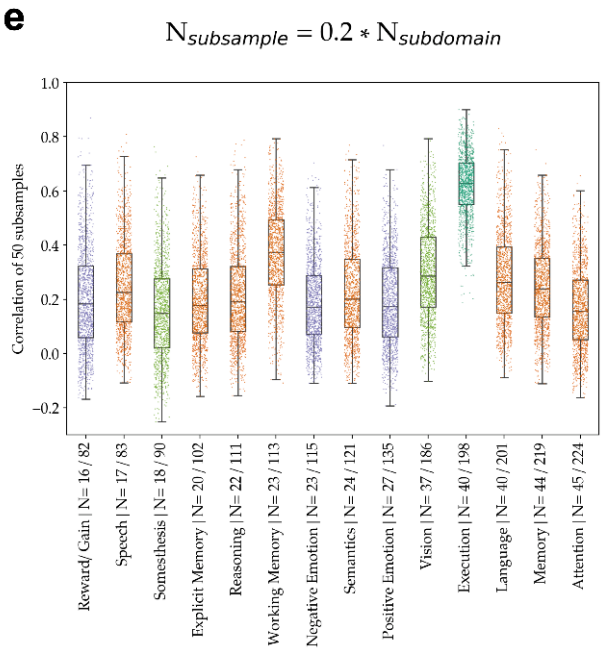

e
$N_{subsample} = 0.2 * N_{subdomain}$
Correlation of 50 subsamples
Reward/ Gain | N= 16 / 82
Speech | N= 17 / 83
Somesthesis | N= 18 / 90
Explicit Memory | N= 20 / 102
Reasoning | N= 22 / 111
Working Memory | N= 23 / 113
Negative Emotion | N= 23 / 115
Semantics | N= 24 / 121
Positive Emotion | N= 27 / 135
Vision | N= 37 / 186
Execution | N= 40 / 198
Language | N= 40 / 201
Memory | N= 44 / 219
Attention | N= 45 / 224

Figure 4: Stability of cerebellum-specific ALE maps. To assess stability of cerebellum-specific ALE (C-SALE) maps, repeated subsampling analyses were performed. Two different subsample strategies were used for both behavioral domains (BDs) and subdomains. One sampled an absolute number of experiments for each subsample (**a, d**) and the other a proportion of overall (sub)domain sample size (**b, e**). For BDs specifically, we created fifty random subsamples of different fixed sizes and proportions (**a, b**). Then, all subsampled unthresholded z-maps within a BD (**a,b**) and subdomain (**d, e**), were correlated. (**a**) Shows spatial correlation between fifty random subsamples at different absolute sizes (n = 25, 50, 75, and 100) within BDs. In turn, (**b**) shows spatial correlations for different proportions of overall BD sample size (proportion (p) = 0.2, 0.4, 0.6, and 0.8). Stability of meta-analytic maps for all BDs increased at higher proportions, visually suggesting stability developed as a logarithmic function. However, stability within Action was higher even at low proportions, and appeared close to plateauing at the .8 proportion. To a lesser extent, stability of Cognition increased faster than other BDs with increasing subsampling proportions. This was especially evident at high (i.e., .8) proportions. (**c**) To highlight stability of thresholded (p_{voxel} < .001 and k = 50) z-maps, percentages of .8-proportion subsamples that reached convergence at any given voxel were mapped to a common flatmap per BD. This illustrated that, whereas unthresholded maps were stable at these proportions, convergence was only reached consistently in Action. Cognition subsamples only sometimes reached convergence in crus I. In (**d, e**) stability of unthresholded maps for subdomains are shown. Here, fifty random subsamples were used as input to C-SALE, consisting either of fifty experiments per subsample (**d**), or of a number proportional to subdomain sample size (p = .2) (**e**). Whereas with increasing subdomain sample size (and thus decreasing proportion) stability decreased, Action Execution stood out as being remarkably stable. (**e**) Serves to compare subdomains. Generally, .2-proportion subdomain and BD (**b**) subsamples were similarly stable. Here, Action Execution was elevated above other subdomains. Notably, stable unthresholded maps (e.g., Action (Execution), Vision, Working Memory, and Language) reached convergence (**Figures 1-3**).

heatmaps report mean z-values for each C-SALE map within each parcel, instead of correlations. For both strategies, the reported heatmaps were hierarchically clustered. Asterisks denote significant spatial correspondence after accounting for spatial autocorrelation (SA) and multiple comparisons ($p_{variogram,\ FDR}$ = .05).

First, MDTB[60] comparisons revealed two main clusters, largely separating Cognition and Emotion from Action and Perception (including subdomains) (**Figure 5a**). Of 840 combinations, sixty-eight correlated significantly. Notably, Social Cognition was spatially correlated with nine of fourteen maps. Math and Landscape Movie (**Figure 5b**) correlated with most C-SALE maps (nine of forty for both), whereas Theory of Mind and Verb Generation correlated with several C-SALE maps involved in Social Cognition and several aspects of Language. For the mid-granularity hierarchical atlas[56] (**Figure 5c**), two main clusters separated Action (Execution) from all other (sub)domains (**Figure 5d**). Action C-SALE maps corresponded significantly to M2 and M3 parcels. Significant correspondence was also found between Action Observation and A1, Explicit Memory and D1, and Social Cognition and S3 and S4. For putative behavioral labels of these parcels, see **Supplementary Figure 11c**. For comparisons with cerebellar resting-state atlases[63,95], lobular definitions[54,92,93], and functional gradients[94], see **Supplementary Results** and **Figures 11-13.** Full correlations or mean

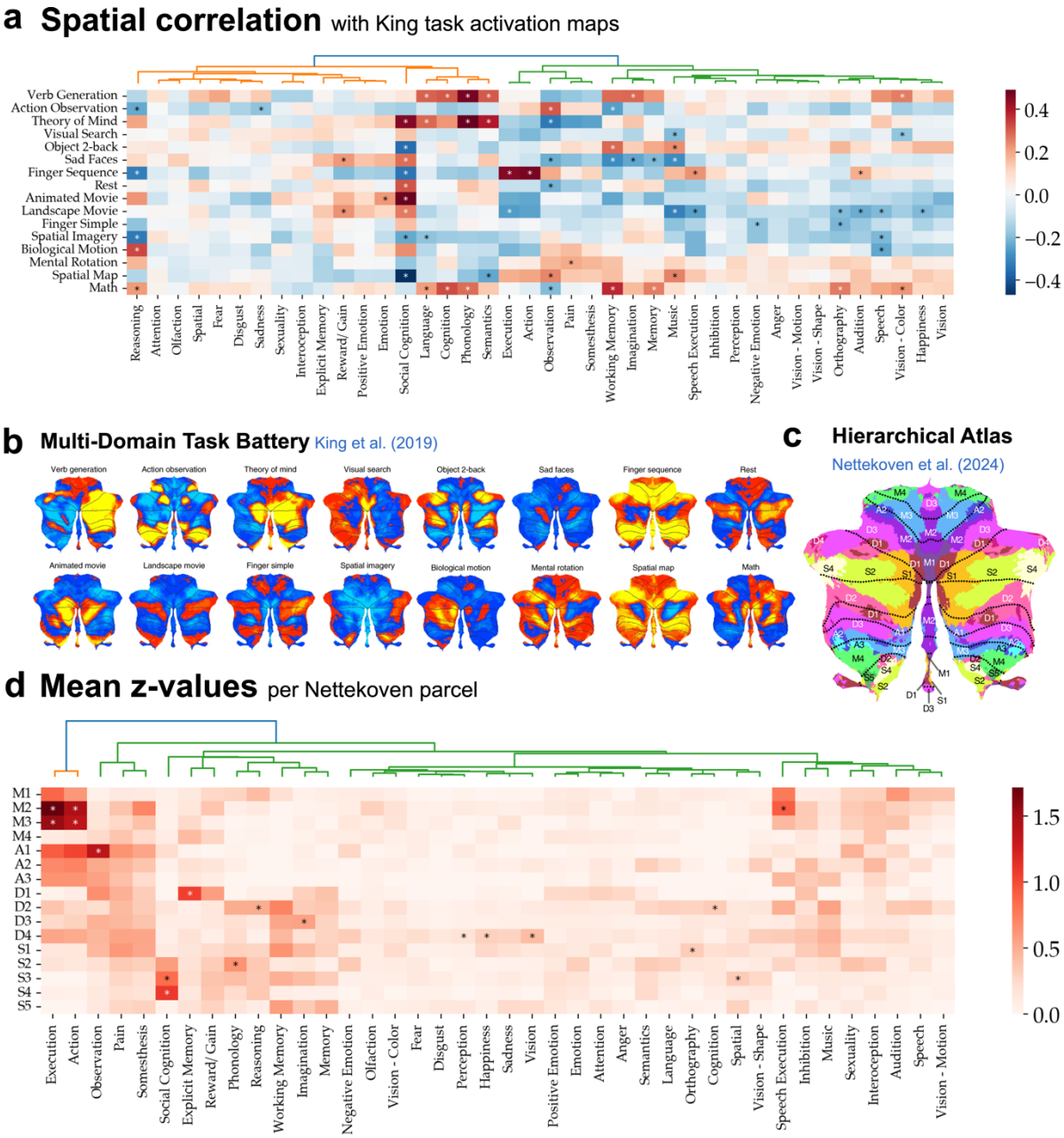


Figure 5: Correspondence of cerebellum-specific ALE maps with King task maps and Nettekoven hierarchical atlas. To contextualize behavioral maps and relate them to popular cerebellar parcellations, we compared unthresholded cerebellum-specific ALE (C-SALE) z-maps with the King multi-domain task battery (MDTB)[60] **(b)** and the symmetrical mid-granularity Nettekoven atlas[56] **(c)** that was merged across left and right parcels. **(a)** Illustrates spatial correlations between each (sub)domain's z-map and group-level task activation map included in the MDTB's publication[60]. Variograms were used to account for spatial autocorrelation (SA). Asterisks indicate significant correlations ($p_{variogram,\ FDR} < .05$). **(d)** Illustrates mean z-values of each (sub)domain's C-SALE map within each parcel. Asterisks indicate significant combinations ($p_{variogram,\ FDR} < .05$). Both heatmaps **(a, d)** were hierarchically clustered. For putative functional labels of the Nettekoven parcels in (**d**), see **Supplementary Figure 11c**. (**b, c**) were adapted from[60], and[56], respectively, after obtaining permission from the authors.

z-values, as well as *p*-values for all comparisons can be found in **Supplementary File 7**.

Meta-analytic connectivity modeling

Recognizing that much of cerebellar functional topography reflects connectivity, we last use C-SALE clusters as seeds for whole-brain MACM analyses, revealing brain-wide coactivation networks. Here, the overall set of analyses was restricted to (sub)domains showing convergence ($p_{voxel} < .001$ and k = 50) in C-SALE analyses (Action and eleven subdomains). Experiments within each (sub)domain were then restricted to those that had at least one coordinate within the regions converging in C-SALE analysis (**Figures 1-3**). To provide stable MACM analyses, an additional prerequisite was that at least seventeen such experiments existed[96]. This further restricted the set of valid analyses to four subdomains (Execution, Execution Speech, Working Memory, and Vision). For each analysis, the overall number of experiments and coordinates are reported in **Supplementary File 8**.

Updates to the null hypothesis

As for C-SALE, instead of assuming spatial homogeneity, we updated the null hypothesis of MACM to reflect the unequal probability distribution of finding foci across the brain (**Supplementary Figure 2**). This ensured that convergence in a behavior was assessed relative to baseline activity. Our implementation differs subtly from a previous implementation of MACM that accounts for unequal distributions of reported effects. Specific CoActivation Likelihood Estimation (SCALE) uses the set of coordinates in the relevant dataset to pseudorandomly sample null coordinates[97]. The current method uses a probabilistic space to sample biased coordinates. This has two primary advantages: 1. it incorporates sample sizes to recognize differential confidence in coordinates; and 2. it provides a smooth, continuous sampling space.

Meta-analytic connectivity maps

Ultimately, five MACM analyses were performed. Thresholded MACM ($p_{voxel} < .001$ and k = 50) coactivation maps for all (sub)domains can be found in **Figure 6**. Unthresholded maps can be found on GitHub (see *"Data Availability"* and *"Code Availability"*). To aid interpretation of cluster locations, common cerebellar[54,92,93], subcortical[98], and cerebral cortical[95,99] parcellations are plotted in **Figure 6a**. MACM results for the Action BD across brain

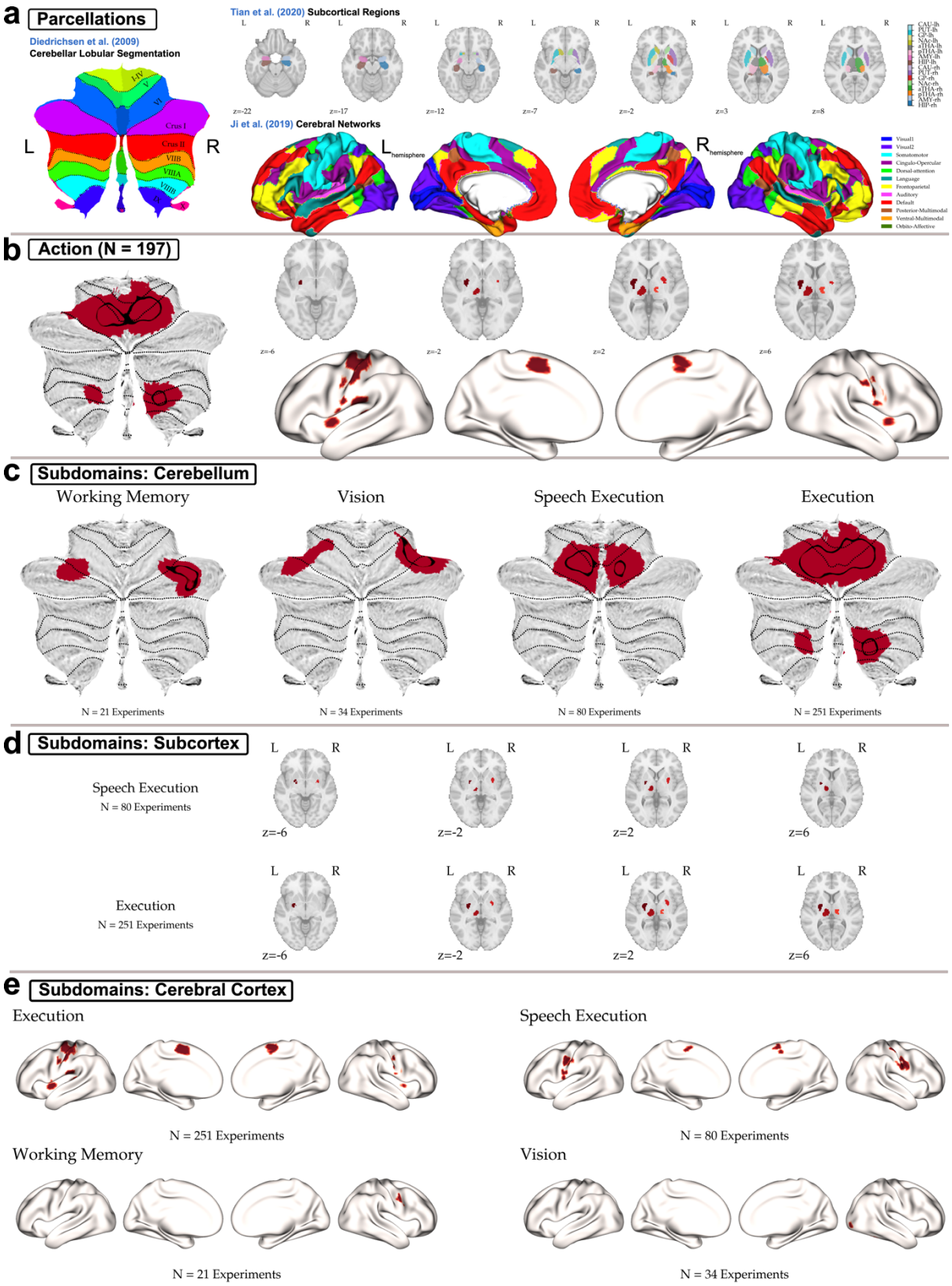


Figure 6: Whole-brain coactivation networks from cerebellum-specific ALE clusters. Whole brain coactivation networks for behavioral domains (BDs) and subdomains. Specifically, the maps show results from meta-analytic connectivity modeling (MACM) using the whole-brain probability distribution for construction of the null models (see **Figure 1**, **Supplementary Figure 2**). Clusters of activity convergence in cerebellum-specific ALE (C-SALE) (**Figures 2, 3**) were used to restrict experiments for each MACM analysis (MACM seeds). Across the figure, black outlines are overlaid on the cerebellar flatmaps to illustrate the seed region. (**a**) Shows cerebellar, subcortical, and cerebral parcellations used to contextualize MACM maps.

These are the lobular atlas for the cerebellum[54,92,93,100], a subcortical regional parcellation[98], and cerebral cortical twelve-network parcellation[95]. (**b**) Shows full MACM maps across cerebellum, subcortex, and cerebrum for the Action BD. (**c-e**) Show MACM maps for subdomains, for cerebellum (**c**), subcortex (**d**), and cerebral cortex (**e**), respectively. In (**d**) only Action Execution and Speech Execution are shown, as they were the only subdomains to reach subcortical convergence. Abbreviations (**a**): rh = right hemisphere; lh = left hemisphere; HIP = hippocampus; AMY = amygdala; pTHA = posterior thalamus; aTHA = anterior thalamus; NAc = nucleus accumbens; GP = globus pallidus; PUT = putamen; CAU = caudate nucleus.

subdivisions are presented together in **Figure 6b**. For subdomains, cerebellar MACM results are reported with black outlines of the seeds (**Supplementary Figure 14**) overlaid (**Figure 6c)**; those for the subcortex on 2D (MNI space[91]) mosaics (**Figure 6d**); and those for the cerebral cortex onto the 32k FSLR surface[101,102] (**Figure 6e**). Cerebellar MACM maps illustrated increased regions of convergence relative to their seeds (**Figure 6b, c)**. Although such patterns are expected, it is noteworthy that this revealed symmetrical coactivating cerebellar regions across Action (Execution) (both in left VIIb-VIIIb), Working Memory (left VI-crus I), and Vision (left V-crus I). Only Action, Execution, and Speech Execution converged in the subcortex (**Figure 6b, d**). All overlapped bilateral caudate and the left nucleus accumbens and thalamus. Whereas Action (Execution) convergence additionally somewhat overlapped the right nucleus accumbens and anterior thalamus, Speech instead overlapped the bilateral putamen (**Supplementary Figure 15**). Each of the (sub)domains converged in the cerebral cortex (**Figure 6b, e**). It is worth noting additional aspects of these maps. First, across Action, Execution, and Speech Execution, MACM maps were highly similar. Secondly, they were highly symmetric across the brain (**Figure 6b-e**). Whereas this also holds true for the cerebellum for Working Memory and Vision (**Figure 6c**), asymmetric clusters can be noted in the cerebral cortex (right for both) (**Figure 6e**). Overall, we showed that cerebellar behavioral clusters are systematically coactivated with distinct brain regions in cerebral cortex and subcortex.

Discussion

In the current study, we created a meta-analytic mapping of the cerebellum across forty behavioral (sub)domains. Specifically, we aimed to refine the cerebellar behavioral topography by using a new meta-analytic method to summarize large-scale behavioral imaging data while accounting for literature reporting biases. This facilitates performing accurate cerebellar meta-analyses, also in data without full cerebellar coverage. We first describe the prominent literature bias in the distribution of reported effects (foci), likely exacerbated by historical cerebellar neglect. To overcome biases in foci, we created a new implementation of ALE,

called C-SALE. We detail how we incorporated reporting biases into an updated null model for both cerebellar and whole-brain meta-analyses. The new method greatly altered locations of convergence in the cerebellum. Notwithstanding, activity across human behaviors (the Action BD and eleven of thirty-five subdomains) converged onto distinct cerebellar subregions. Converging subdomains covered aspects of Action, Cognition, Emotion, and Perception. Subdomains belonging to each BD occupied distinct areas of the cerebellum: anterior and posterior motor representations were broken up by cognitive-emotional mappings in posterior-lateral lobes. Four Perception subdomains reached convergence in lobules V-crus I. However, most behavioral (sub)domains did not reach convergence despite illustrating distinct patterns of subthreshold convergence. Repeated subsampling revealed that behaviors mapped to subregions moderately stably across random subsets of experiments. Cognition, Perception, and especially Action (sub)domains stood out for having relatively consistent localizations. Connectedly, confidence in these stable localizations increased, and thresholds were more readily reached. Spatial correspondence to parcellations and mappings[56,60,63,86,94,95] revealed that converging clusters often mapped relatively strongly to one or several existing parcels or maps. Locations of subthreshold convergence also often colocalized significantly with (behaviorally) related parcels and maps. Lastly, MACM was performed to reveal coactivation networks for several aspects of Action, as well as Working Memory and Vision.

Methodological implications of C-SALE for cerebellar mapping

Classic ALE was unsuitable for cerebellar CBMA, since its null assumption (equally distributed foci)[66] was violated. The unequal distribution of foci was likely exacerbated by incomplete cerebellar coverage. Previous cerebellar CBMAs have often – but not always – used cerebellar coverage as inclusion criterion. Thus, convergence may be overreported in superior regions (and vice versa), misassigning behaviors to cerebellar subregions. Given the probability hotspots, convergence in regions V-crus I may be most at risk of being overreported. For future CBMAs, we recommend accounting for unequal distributions of reported effects alongside explicitly verifying full (cerebellar) coverage. Although manually curated CBMAs often do the latter, the former is important to account for brain-wide heterogeneity in reporting patterns. Our method can be flexibly adapted to analyses at the whole-brain-level or any brain ROI.

Recently, a similar cerebellar meta-analytical study was published[86]. Here, CBMAs across the NeuroSynth database[87] resulted in a behaviorally relevant cerebellar parcellation. Whereas

NeuroSynth contains more articles (14,371 versus 3,406), BrainMap is manually indexed and includes information on sample size[88], allowing assigning differential confidence to foci. Notably, the NeuroSynth study also reports a severe superior-inferior bias in reported effects, leading to exaggerated superior convergence. The authors discuss how their winner-takes-all strategy, comparing ALE scores at the voxel-level, protects against spatial biases[86]. Though this is an important step in cerebellar behavioral mapping, ALE scores still reflect a combination of true biological signals, different sampling efforts, sample sizes, and experimental consistency, alongside the biased baseline of reported effects. Our results, especially the subsampling analyses, suggest that each of these factors may play a – largely unknown – role in producing ALE maps. Hence, comparing ALE values, even at the same spatial location, will always favor some behaviors/ behavioral domains. Secondly, the underlying assumption that no biased and systematic relation between neglect of the cerebellum and task domains exists[86] needs to be examined. Many have pointed out how the cerebellum has historically been neglected in non-motor functions (see primarily the consensus paper series[7–13]). We briefly explored this notion, mapping reported effect locations versus cerebellar voxels for each BD separately (**Supplementary Figure 16**). Although conflated by actual behavioral signal, increased inferior effects within Action suggests that full cerebellar coverage may be more common in this BD. This, however speculative, underlines the need to formally investigate biased cerebellar neglect and not assume *a priori* its absence (nor existence).

C-SALE refines the cerebellar behavioral topography

Twelve of forty behavioral (sub)domains reached significant convergence in specific cerebellar subregions. Importantly, this does not imply heightened regional activity relative to the whole brain but to the cerebellum. Briefly, our localizations support cerebellar subregional roles across a great diversity of behaviors, which supports a growing consensus[7,8,10,11]. Specifically, we report highly stable Action localizations (including Execution and Speech Execution) consistent with a dual motor mapping[63,65,103]. Additionally , these motor representations are broken up by bilateral cerebellar subregions dominantly involved in Emotion and Cognition[65,78–80,82,104]. These cerebellar regions, primarily crura I-II, have received considerable attention due to their role in integrative, transmodal networks across the brain[18,24,105,106]. These lobules are characterized by massive primate-general expansion[25], perhaps exceeding that of the prefrontal cortex[107], underlying their relevance for sophisticated

primate behavior. Lastly, we report converge within the Perception BD. We interpret the locations of convergence in cerebellar subregions relative to the whole cerebellum more exhaustively in the **Supplementary Discussion**. Note that we also discuss interesting patterns of subthreshold convergence, which should be interpreted as inconclusive speculation. In converging regions, we can be confident of the localization. For non-converging regions, more high-quality data will be necessary to assign behaviors confidently.

Stability of cerebellar mapping differs across behavioral domains

Next, we assessed stability of C-SALE maps. Action was highly stable: parts of the anterior representation reached convergence across all, and the posterior across most, subsamples. Cerebellar motor areas are stably, reciprocally connected to the basial ganglia and cerebral motor cortex across primates[9,15,17,20–23,49,50]. Such consistent connectivity may lead to consistent functional localizations[18,41]. Cerebellar parcellations have repeatedly placed motor representations in similar locations[56,60,63,95], as have task-based localizations[65] and precision-mapping approaches[55,64]. Action (Execution) MACM corroborated consistency of the brain-wide motor network[63,95]. Likewise, somewhat elevated consistency of Language (several aspects), and (Working) Memory, may be supported by connectivity. Language and memory are examples of functions with well-defined extracerebellar neural substrates. The cerebellum has extensive reciprocal connections with prefrontal cortex, thalamus, and cerebral language areas[4,8,14–16,20,21,76,108]. Structural connections with hippocampus[109] and amygdala[110] support cerebellar involvement in socio-effective shaping of motor-related behaviors[111], and normal hippocampal functioning including memory and navigation[109,112].

Cerebellar functional connectivity can be organized along a unimodal-transmodal axis[94,113,114] (as can cerebral[115,116] and whole brain[113] connectivity). Gradients of functional abstraction[117] transcriptomic and molecular expression[118,119], and granule cell physiology[120] underline gradual organizational aspects of the cerebellar cortex. Relative to unimodal connectivity patterns of the motor brain network, cerebellar afferents in other behaviors may thus be more transmodal, eliciting activity across smaller, distributed cerebellar areas[121]. Each set of afferents may be part of relatively separate reciprocal networks involved in distinct functions (see: cerebellar modules[122–124]). Even adjacent modules can be involved in different functions, making it difficult to expose their functions using CBMAs. Emotion and Interoception, with low subregional preference, are good examples of domains where overcoming the issues of partial volume effects in the (small, folded, often misaligned[54,100]) cerebellum are especially

important. Even if these task domains elicit consistently elevated activity across small distributed or interdigitated cerebellar areas, the summarizing nature of CBMAs combined with these cerebellar challenges may preclude finding statistical convergence. Importantly, greater experimental consistency in e.g., Action, Language, Working Memory, and Vision may also lead to increased meta-analytic stability. Whereas there is a finite number of ways to perform motor, perceptive, and some cognitive behaviors, emotional and interoceptive behaviors may be more diverse, as may the ways to elicit and measure them. Together, low stability and differences between C-SALE mappings and previous CBMA clusters warn for caution in interpretation of small-to-intermediate-sized CBMAs.

Limitations of the current study

Several limitations of this study are important to consider. First, we could not monitor experimental consistency within (sub)domains. Each analysis may include experiments probing subtly different behavioral aspects and extents of overlap with other behaviors. Therefore, we could not fully disentangle experimental conditions from cerebellar organization. Typically, CBMAs start with a literature search, followed by manual text scanning to homogenize experiments[88,125,126]. Since this was not feasible, we focused on data-driven interpretation of the largest possible sample. Consequently, every result is an opportunity for more specific research questions aiming to map behaviors more precisely. Experimental sample sizes per (sub)domain may increase after manual curation, as we combined all experiments within a study to – perhaps conservatively – prevent experimental overlap[125].

Next, decisively mapping behaviors to cerebellar subregions was no trivial task. Incongruencies with previous parcellations and mappings, and a lack of subsample stability show that complementary perspectives are necessary to understand cerebellar behavioral topography. We report many correspondences between C-SALE maps and behaviorally related aspects of mappings and atlases[56,60,63,86,95], and between MACM maps and subcortical parcellations and cerebral networks[95,98,99]. However, many behavioral (sub)domains mapped moderately or not at all to previous parcellations. This implies that these (sub)domains may be poorly represented by current state-of-the-art parcellations. Importantly, the discriminatory performance of their borders makes parcellations useful for many applications. However, our findings suggest there is need for a large-scale behavioral cerebellar topography. It would ideally be created directly from full task-activation maps (not only peaks), span task domains,

and include many – diverse – individuals. Ultimately, substituting coordinate-based for map-based meta-analyses (while still accounting for biases) may be the way forward.

For the present study, low resolutions of older data in meta-analytic databases may ultimately limit accuracy of localizations. Scanning at 7T can greatly improve cerebellar resolutions, as can using cerebellum-optimized sequences[127] and dielectric pads[128]. Although we may be close to limits of CBMAs in the cerebellum with current data, increased resolutions will improve localizations. High-resolution cerebellar data[128,129] in many individuals can also help improve cerebellar alignments[54], essential given the summarizing premise of CBMAs.

Conclusion

Here, we used the BrainMap database and an adaptation of the ALE method (C-SALE) to perform large-scale CBMAs across forty behavioral (sub)domains. Our findings underline the systematic omission of the inferior cerebellum in neuroimaging data. C-SALE overcomes these biases, improving accuracy of CBMAs even without full cerebellar coverage. We show that behaviors across Action, Cognition, Emotion, Interoception, and Perception converge onto distinct cerebellar subregions. Our maps refine functional subdivisions from previous mappings and parcellations. We also show that behaviors within Cognition, Perception, and especially Action (motor execution) map more systematically to cerebellar subregions, which may be related to cerebellar connectivity profiles and experimental consistency. In sum, we reveal a comprehensive meta-analytic topography of the cerebellum that implicates cerebellar subregional involvement across human behaviors. However, our new method also suggests that caution in interpretation of previous localizations, which may have been liberal or inaccurate, is warranted. To improve behavioral localizations in the future, we make available our methods for cerebellar or any other brain volumetric ROI. In closing, subareas of the cerebellum consistently activated in distinct behaviors offer putative regions of interest for functional imaging, neurostimulation, and ultimately diagnosis and intervention.

Methods

Overview

In this study, we leveraged the BrainMap database[88] to pull the largest available sample of manually indexed fMRI and PET literature at the whole-brain-level. Comprehensive access to BrainMap data and metadata was authorized by a collaborative use agreement (brainmap.org/collaborations). From the database, we initially obtained data from 8,408 whole-brain experiments (2,214 studies; 32,836 participants). Of these, 2,415 experiments (1,129 studies; 16,410 participants) reported peak coordinates within the expanded cerebellar mask. After merging experiments within each study[125], 1,129 unique experiments remained. From these, we constructed C-SALE maps across five BDs (Action, Cognition, Emotion, Interoception, and Perception) and thirty-five more fine-grained behavioral subdomains. These include several aspects of language and memory, fear and reward, and most of the outward senses. To recognize cerebellar clusters in relation to the whole-brain, we subsequently performed MACM. Findings at the cerebellar and whole-brain level were compared to existing parcellations. For C-SALE results specifically, we also assessed stability of cerebellar localizations across behavioral (sub)domains.

Study Inclusion Strategy

Experiments were included in C-SALE analyses based on the following criteria: **1.** To construct a functional map of the general healthy population, only normal mapping studies that measured within-subject activation contrasts in healthy controls (no interventions; number of subjects eight to forty-four) using fMRI or PET were included **2.** To localize behaviors to cerebellar subareas, only experiments with a peak coordinate within the cerebellar ROI were included. The cerebellar ROI was isolated using the 1 mm resolution cerebellar template[54], dilated 6 mm in each direction (hereafter "dilated cerebellar mask") to account for spatial misalignment across studies. This step also aimed to overcome cerebral cortical BOLD-bleeding. **3a.** Separate datasets were created for the five BDs indexed in BrainMap. Only activation contrasts were considered, as deactivations often had insufficient numbers of experiments. Activation foci were limited to those within the mask in **2**. and those included within each BD. **3b.** Next, in the same way, datasets were created for subdomains indexed under the five BDs. **4.** To prevent unjustly embellishing statistical power, it is important to prevent overlaps in included experiments[125]. Within the BrainMap database, single studies may contain multiple experiments with potential overlaps in experimental contrast and participants. Hence, we merged the coordinates within a study to represent a single experiment. This validated

approach[130] minimizes within-group and within-experiment effects[131], preventing overlapping experimental contrasts excessively contributing to convergence[125]. As empirical simulation indicates that seventeen or more experiments should be included for stable meta-analysis[96] ultimately five eligible BDs and thirty-five eligible subdomains ($N_{experiments} \geq 17$) datasets were created.

Construction of behavioral cerebellar maps with Cerebellum-Specific ALE

After collecting datasets for each sub(domain), we performed ALE for each using an in-house modification of the Neuroimaging Meta-Analysis Research Environment (NiMARE) version 0.2.0rc3[132,133]. NiMARE is a Python package that facilitates programmatic interaction with BrainMap activation foci and performing CBMA based on, among other algorithms, ALE[66,67]. We added an efficient implementation of ALE using graphical processing units (GPUs) (github.com/amnsbr/nimare-gpu). This facilitated running the calculations of many permutations of ALE, and the multiple experiments therein, in parallel. These permutations are necessary to generate null samples for each meta-analysis, as described below. In the GPU implementation, calculations pertaining to individual permutations, experiments, and foci were parallelized at two levels, across GPU "blocks" (permutations and experiments) and "threads" (foci). Ultimately, parallelization was highly advantageous because of the high number of performed meta-analyses across different sub(domains) and subsample configurations. The GPU implementation led to considerable speed-ups of calculations when scaling the number of permutations and experiments. For example, we observed a speed-up of 101.3x in GPU versus central processing units (CPUs) when running a rather typical SALE analysis for 100 experiments with 10,000 permutations (**Supplementary Figure 17**).

Essentially, ALE tests the distribution of experimental activation foci against a null distribution that assumes random spatial associations across cerebellar GM. Put simply, it finds spatial locations where activations converge more than chance. However, visualizing foci, we observed strong spatial biases in reported effects (**Figure 1a, Supplementary Figure 1)**. These biases renders the standard null hypothesis (**Figure 1b)** unsuitable[66,67]. Hence, we aimed to account for biases by incorporating the distribution of foci into the null model against which datasets were tested. Our approach is comparable to the SCALE approach used for MACM[97]. Since we here adapt this approach to the cerebellum, we refer to it as cerebellum-specific activation likelihood estimation (C-SALE). Note that contrary to this name, the method can be flexibly adapted to any brain volume. To assess improvements of the new method, we also ran

BD-level analyses with standard ALE. Its specific methods are discussed here first. For full illustration of the commonalities and differences between classic ALE and C-SALE, see **Figure 7**.

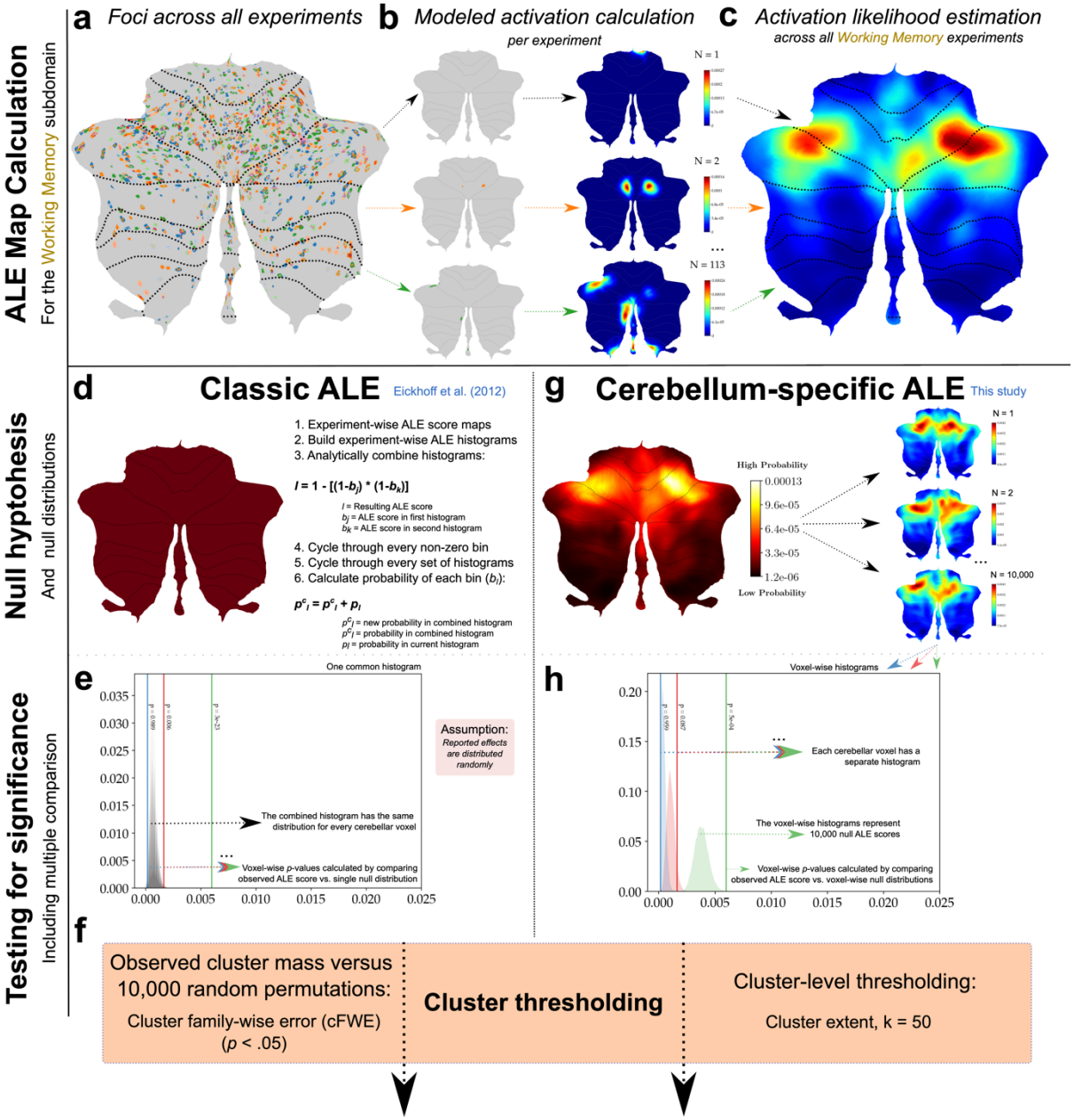


Figure 7: Commonalities and differences across classic ALE and cerebellum-specific ALE methods. Full pipelines for both the classic implementation of ALE and the new method, cerebellum-specific ALE (C-SALE). (**a-c**) ALE score map calculation is the same across classic ALE and C-SALE. (**a**) First, peak activation coordinates (foci) across a task or behavioral domain are collected. Data from the Working Memory subdomain is used as example here. (**b**) Coordinates are then convolved by a gaussian kernel with a full width half maximum inversely proportional to experimental sample size. The convolved foci maps are combined into modeled activation (MA) maps by taking their voxel-wise union. (**c**) Last, MA maps are combined into an ALE score map, again by taking the voxel-wise union across all (here 113) maps. Classic ALE and C-SALE start diverging due to different null hypotheses (**d, g**). We first discuss classic ALE (**d-f**). (**d**) Classic ALE relies on the assumption of spatially randomly distributed reported effects, as illustrated by a flat cerebellar sampling space. The ALE score map

calculated in (**a-c**) is tested against this null hypothesis to reveal behavioral activations that go beyond truly random spatial effects. The current implementation of ALE[67] bypasses the (costly) need for permutation, by analytically combining experiment-wise ALE score histograms (thus: at the MA map-level (**b**)). This analytical strategy is outlined in (**d**). (**e**) Ultimately, this leads to one common histogram of ALE scores, against which each voxel's ALE score (as determined in (**a-c**)) is compared to obtain *p*-values. (**f**) A permutation-based strategy (cluster family-wise error; cFWE) is used to determine how observed cluster mass compares to cluster masses in 10,000 randomly generated maps. Clusters are considered significant if they exceed $p_{cFWE} < .05$. Next, we discuss C-SALE (**g, h, f**). (**g**) In C-SALE, instead of assuming spatial homogeneity of reported effects, unequal spatial distributions are assumed. Specifically, a null probability distribution is created by convolving all brain-wide experiments in the BrainMap database and taking the voxel-wise sum of their probabilities. This probabilistic map is then normalized to one within the cerebellum, describing the base likelihood of finding foci at cerebellar voxels (**g**; left). To determine whether the observed voxel-wise ALE scores converge beyond this baseline, a permutation-based strategy is used. Specifically, 10,000 MA null maps are created by pseudorandomly sampling coordinates relative to the probability distribution. Each of these MA maps is biased by the baseline activation rates (**g**; right) and this is reflected in shifts in ALE score histograms across voxels (**h**) (three random voxels are illustrated here). (**h**) This means that when testing the ALE score map of a behavioral domain (**a-c**), each cerebellar voxel is tested against its unique distribution of ALE values. As illustrated by the green histogram, voxels that are reported more will require higher ALE scores to reach significant convergence ($p_{voxel} < .001$). (**f**) Lastly, to threshold cluster sizes, we opted for a lower limit of k = 50 voxels to prevent very small cerebellar regions reaching convergence. In the future, a (non-trivial) adaptation of cFWE for our probabilistic method may further improve cluster-wise inference.

Classic ALE

The classic implementation of ALE[66,67] assesses convergence of reported effects (activation foci) against a null hypothesis that assumes foci are distributed equally across the cerebellar GM. For each analysis, we did the following: **1.** Convolve activation foci (**Figure 7a**) by a 3D Gaussian kernel. Importantly, these kernels had a full-width half-maximum (FWHM) inversely related to sample size. This means narrower distributions – and thus higher statistical certainty – were assigned to experiments with large samples (and vice versa). **2.** The convolved foci were then combined, first at the level of individual experiments. The union of all convolved foci within an experiment was taken, creating modeled activation (MA) maps (**Figure 7b**). These maps reflect voxel-wise maxima across convolved foci. **3.** Next, experiments within a BD were combined. This was again done by taking their union, resulting in an ALE score map for each BD (**Figure 7c**). **4.** These ALE score maps were then tested against the null assumption of random spatial associations (**Figure 7d**)[66,67], resulting in voxel-wise *p*-values (**Figure 7e**). **5.** These *p*-values were thresholded at $p_{voxel} < .001$. Using a permutation method, 10,000 random sets of experimental foci were created under the null hypothesis. For each of these coordinate sets, steps **1-5** were repeated. **6**. Lastly, cluster-level family-wise error (cFWE) was used to test statistical significance of resulting clusters from step **5** while accounting for multiple comparisons[67]. Here, cluster extent (mass) of the actual experimental set in each

sub(domain) was compared against the distribution of cluster extents in the 10,000 random ALE score maps. Cluster extents were thresholded at $p_{cFWE} < .05$ (**Figure 7f**).

Accounting for the spatial bias of reported effects in probabilistic C-SALE

However, as discussed above, there were spatial biases, with many more foci in superior regions. This rendered the standard null hypothesis unsuitable (**Figures 1, 7**). In C-SALE, we tested where reported effects converged in cerebellar subregions beyond the baseline, spatially biased, probability (**Figures 1b, 7g**). We outline our method for calculating voxel-wise baseline (null) probabilities in reported effects. Thereafter, we describe how we tested each (sub)domain's ALE score map against this new null hypothesis.

Creating a bias-accounting null model

To calculate voxel-wise baseline probabilities within the cerebellum, we first created a whole-brain probability map by taking all activation foci in the BrainMap database (its February 2024 release) for healthy adult contrasts (8,408 unique experiments; 2,214 studies; 32,485 participants; 69,703 coordinates). Then, each coordinate was convolved by a 3D Gaussian kernel with FWHM inversely proportional to experimental sample size. Next, convolved foci were summed voxel-wise to obtain a whole-brain probability map of reported effects. The resulting map was subsequently masked to the dilated cerebellar mask and normalized to a sum of one. This resulted in a voxel-wise baseline probability map of finding foci at any cerebellar voxel, regardless of task domain. Constructing a probabilistic, rather than deterministic[97], null model has some advantages. By convolving foci by a FWHM based on experimental sample size (included in BrainMap[88]), the null model accounts for foci uncertainty. Connectedly, even regions without foci can be sampled into the null model at low probabilities, creating a continuous sampling space. By using the dilated cerebellar mask, the probabilistic null model may also partly account for influences of BOLD-signals from adjacent visual and temporal cerebral cortices. Unequal distributions of reported effects are ubiquitous across the brain (**Supplementary Figure 2**). Hence, by calculating the sum of convolved foci at the brain-level and then restricting this probability distribution to the cerebellum, the null model can be flexibly adapted to any volumetric ROI. Note that the null distribution is dataset-specific and needs to be adapted for different or updated datasets.

Creating meta-analytic maps with C-SALE

Following this calculation of the baseline (null) probability map, we performed C-SALE for each (sub)domain as follows: **1.** The ALE score maps were calculated in the same way as for the classic ALE approach (**Figure 7a-c**). **2.** In contrast to classic ALE, here we constructed voxel-wise null distributions for each (sub)domain separately using Monte Carlo permutations (n = 10,000). Specifically, in each permutation, pseudorandom foci were sampled from the dilated cerebellar mask, weighted by the baseline probability map (**Figure 7g**). Next, a null ALE score map was calculated. **3.** Each voxel's observed ALE score was tested against its specific null distribution to calculate voxel-wise *p*-values (**Figure 7h**). **4.** The resulting *p*-values were thresholded at $p_{voxel} < .001$. Cluster extent was thresholded at k > 50 (**Figure 7f**). Note that adapting the preferred method for cluster-extent thresholding, cFWE, is not trivial due to the probabilistic nature of our method. Future work may improve cluster-extent thresholding for the C-SALE method.

Comparing performance of C-SALE to classic ALE

We assessed if C-SALE showed improvements in specificity of cerebellar clusters. The main aim of our study was to identify cerebellar subregions with above-chance convergence across behaviors, relative to the cerebellum. Thus, first, we could compare z-maps resulting from both methods and inspect cerebellar coverage across domains (**Figure 7c, d**). Secondly, we used spatial correlations across (sub)domains as a proxy for specificity. Whereas some spatial overlap is expected, extensive overlap implies the method was not able to differentiate (sub)domains (low specificity). For every pair of (sub)domains, we calculated spatial correlation (Pearson's R) of unthresholded z-maps and assessed the distribution of correlation coefficients (**Figure 7e, f; Supplementary Figure 3**). We also compared spatial correlations of unthresholded z-maps for all pairs of subdomains (**Supplementary Figure 4**).

Assessing stability of cerebellar C-SALE maps

After establishing that C-SALE provided substantial improvements over classic ALE, we investigated C-SALE maps through a repeated subsampling strategy. (Sub)domains contained an inconsistent number of experiments, and convergence is more likely in larger analyses. Hence, we wanted to assess the stability of different (sub)domain's C-SALE maps (see **Figures 1-3)** considering sample size differences.

To give complementary perspectives on the (sub)domain maps we used two subsampling strategies: 1) with subsample sizes fixed at fifty experiments; and 2) with subsample sizes proportional to the (sub)domain ($n_{subsample} = .2 * n_{(sub)domain}$). First, sampling fifty experiments per subsample examined the effect of using the same absolute number of experiments, since convergence is more common in larger datasets. However, here the proportion of experiments (from the total sample) is vastly different when examining rather intermediate (e.g., Interoception (n = 70) to large (n = 750) samples. Therefore, we also constructed maps at a fixed .2-proportion, providing an additional perspective on stability that facilitates fairer comparison across (sub)domains. These subsample analyses were run for five BDs and the seventeen subdomains with $n_{experiments} \geq 100$. For each (sub)domain and subsampling strategy, we created fifty random sets of experiments from the total set of experiments, rerunning C-SALE in each. We then calculated the distribution of spatial correlation coefficients (Pearson) across pairs of unthresholded subsampled z-maps within each (sub)domain. To understand how stability of unthresholded z-maps developed more comprehensively, we performed an additional set of subsampling analyses for BDs only. Here, we created subsamples at different absolute (n = 25, 50, 75, 100) and proportional (.2; .4; .6; and .8) sample sizes. Again, for each configuration, fifty random sets of experiments were selected. C-SALE was rerun in every set, after which we assessed the spatial correlation across the fifty resulting maps (for each combination of BD and sample size or proportion).

To visualize stability of converging results specifically, we mapped, per (sub)domain, the proportions of subsamples that reached the threshold for significance ($p_{voxel} < .001$ and k = 50) at each voxel to a common flatmap. This was done for each configuration described above.

Correspondence with previous functional parcellations and mappings

To relate findings to established functional cerebellar subdivisions, we calculated spatial correspondence of our C-SALE maps with several published cerebellar maps. Comparisons were made between unthresholded z-maps and continuous parcellations or mappings, and between unthresholded z-maps and dichotomous parcellations. For both strategies, heatmaps were hierarchically clustered. Asterisks denote significant spatial correspondence after accounting for SA and multiple comparisons ($p_{variogram, FDR} = .05$).

For continuous comparisons, spatial correlations were calculated between every *C-SALE map* and *target map*. BD and subdomain z-maps were compared to the MDTB[60] and probabilistic cerebellar lobular segmentation[54,92,93,100]. Specifically, for MDTB, we calculated spatial

correlation between unthresholded z-maps and the continuous group-level task activation maps. For the probabilistic lobular segmentation[54], C-SALE maps were compared against every lobule separately. In both sets of comparisons, to account for SA, we calculated the actual spatial correlation between unthresholded C-SALE z-maps and each continuous target map parcel, and compared it against the null distribution obtained from correlations with 10,000 SA-preserving surrogate maps (from C-SALE maps) created using BrainSmash[134,135]. SA is inherent in neuroimaging data, as close regions are generally more likely to have correlated activity[134–136]. It is important to account for these patterns when testing for significance, which may otherwise largely reflect spatial proximity.

For dichotomous parcellations, heatmaps report average z-values for each C-SALE map within each (target) parcel, instead of correlations. We assessed correspondence with several atlases: **1.** The Nettekoven hierarchical cerebellar atlas (mid-granularity symmetrical atlas with sixteen parcels after merging left and right hemispheres)[56]; **2.** the Buckner resting-state seven-network cerebellar atlas[63]; **3.** the Cole-Anticevic cortical-subcortical atlas (ten cerebellar parcels)[95]; and lastly **4.** the deterministic cerebellar lobular segmentation[54,92,93,100]. Specifically, for each parcellation, we calculated the average z-values of each (sub)domains unthresholded z-maps within each of the parcels. To assess statistical significance, we used a non-parametric test in which observed mean z-values within each parcel of the original C-SALE map were compared against a null distribution of average z-values in SA-preserving surrogate maps created using BrainSmash[134,135]. We reused the set of SA-preserving surrogate maps used in the continuous comparisons.

Lastly, we calculated loading of each voxel within the thresholded C-SALE mask onto the primary (G1) and secondary (G2) functional cerebellar gradients[94] using the LittleBrain Toolbox[137]. These gradients use orthogonal axes to explain variation in resting-state fMRI patterns. The first two axes, those explaining most variance, separate motor from DMN regions (low-to-high loading on G1) and task unfocused-to-focused regions (low-to-high on G2)[94]. Every voxel is colored by the seven-parcel Buckner resting-state network[63] it colocalized with. Note that sign and unit of gradients are arbitrary.

Meta-analytic connectivity mapping

The functions of cerebellar subareas are tightly connected to those of connected brain regions[18,24,39,41]. Hence, we wanted to examine meta-analytic connectivity profiles across

(sub)domains. We adapted the SCALE method, itself an adaptation of MACM that accounts for baseline activations by sampling null coordinates from the dataset in question[97].

Our method tests activations per (sub)domain not against a deterministic null model, but a whole-brain probabilistic null model (see *Accounting for […] probabilistic C-SALE*). Specifically, for each (sub)domain, we used the thresholded ($p_{voxel} < .001$ and k = 50) C-SALE clusters as seed. We first restricted the set of experiments to those that reported at least one peak coordinate within each 3D seed mask. The set of coordinates included in these experiments were then used as input to whole-brain probabilistic SCALE analyses. These analyses aimed to reveal where in the brain coactivation with the seed regions occurred more than chance, given the biased baseline probability distribution across the brain. Importantly, the set and number of experiments was different from the C-SALE analysis: ultimately, we were able to run analyses for the Action BD and four subdomains ($N_{experiments,\ seed} \geq 17$).

For our probabilistic implementation of SCALE, the method mirrored that described for C-SALE. The only differences were the set of input experiments (limited to the C-SALE mask) and the ROI (the whole brain GM mask; >10 % GM probability at 2 mm resolution (Grey10)). Accordingly, for probabilistic SCALE, the null probability map was constructed by normalizing the sum of MA maps of all BrainMap experiments to one within the whole-brain GM mask (Grey10). Again, for each analysis, we created 10,000 null ALE maps by pseudorandomly sampling foci from the GM mask, weighted by the baseline reporting probability (**Supplementary Figure 2**). Subsequently, in each analysis, the observed ALE score was compared against the permutation-based null distribution to calculate voxel-wise *p*-values. *P*-values were then thresholded at $p_{voxel} < .001$ and k = 50. This revealed where brain-wide coactivations with each (sub)domain's C-SALE cluster occurred more than chance given baseline activations across the brain.

Correspondence of meta-analytic connectivity maps to subcortical-cerebral parcellations

For every MACM result, we calculated spatial overlap with subcortical and cerebral cortical parcellations. This was done for subcortical regions[98], cerebral cortical-subcortical networks[95], and, to interpret findings in terms of brain cytoarchitecture, microstructural cortical types[99,138]. For all parcellations separately, we calculated proportions of thresholded MACM maps that colocalized with each parcel in MNI (2mm) space. These proportions are reported as hierarchically clustered heatmaps.

Data availability

The BrainMap database is publicly accessible online, via the Sleuth filtered-search application (brainmap.org/software), via the BrainMap Community Portal (portal.brainmap.org), or by comprehensive download when authorized by a collaborative use agreement (brainmap.org/collaborations). For this study, we obtained the February 2024 data release and uploaded these data to GitHub (github.com/NevMagi/cerebellum_specific_ALE). Full intermediate data resulting from this publication are also uploaded there, including figure panels. Final figures were put together in Inkscape and can be obtained for reuse upon request.

Code availability

All programmatic code used to obtain the results in this article are made available on GitHub: (github.com/NevMagi/cerebellum_specific_ALE). All code necessary to perform bias-accounting coordinated-based meta-analyses, for the whole brain or any volumetric brain region-of-interest, are made available. Also made available is our graphical processing unit implementation of NiMARE which helps speed up MA map calculation (github.com/amnsbr/nimare-gpu), and works for both classic ALE, and deterministic and probabilistic versions of C-SALE (as in the current study).

Materials and Correspondence

Correspondence can be addressed to Neville Magielse (n.magielse@fz-juelich.de) and Sofie L. Valk (s.valk@fz-juelich.de). Permanent address: Max Planck Institute for Human Cognitive and Brain Sciences, Dr. Sofie L. Valk, Stephanstraße 1A, 04103 Leipzig, Saxony, Germany. Telephone: +49 341 9940-2658 | Fax: +49 341 9940-104.

Ethics statement

This study uses data from the publicly available BrainMap meta-analytic database. Data consists of aggregated fMRI and PET data. Specifically, only peak coordinate locations and sample size were used (and available), and hence no individual participants could be identified. No contact was made (nor possible) with any participants, nor was any individual's data handled at any point.

Acknowledgements

We would like to thank Lennart Frahm for insightful discussions on the activation likelihood estimation method and statistics.

Author contributions

NM. Conceptualized the manuscript; assisted in code writing and analysis; interpreted results; created figures; wrote the manuscript; incorporated manuscript reviews. **AM.** Assisted in manuscript conceptualization; provided manuscript reviews (several occasions); assisted in result interpretation. **SBE.** Assisted in manuscript conceptualization; provided technical support; provided revision comments manuscript. **PTF.** Provided the BrainMap data; provided manuscript review. **AS[#].** Conceptualized the manuscript, wrote code, performed chief analysis, assisted in manuscript writing, provided manuscript reviews. **SLV[#].** Assisted in manuscript conceptualization; provided manuscript reviews (several occasions); assisted in data interpretation; provided supervision.

[#] *Authors **AS** and **SLV** contributed to this manuscript equally and hence share last authorship.*

Funding

NM, SBE, AS, and **SLV** were funded by the Helmholtz Association's Initiative and Networking Fund under the Helmholtz International Lab grant agreement InterLabs-0015, and the Canada First Research Excellence Fund (CFREF Competition 2, 2015–2016) awarded to the Healthy Brains, Healthy Lives initiative at McGill University, through the Helmholtz International BigBrain Analytics and Learning Laboratory (HIBALL). **AM** was funded by the German Academic Scholarship Foundation (*Studienstiftung des deutschen Volkes*). **AM, AS** and **SVL** were additionally funded by the Max Planck Society (Otto Hahn award). **PTF** and the BrainMap Project were funded by the National Institutes of Health (MH074457) and by the Malcolm Jones Professorship of Radiology from UT Health San Antonio.

Competing interests

The authors declare no competing interest.

References

1. Eccles, J. C., Ito, M. & Szentágothai, J. *The Cerebellum as a Neuronal Machine*. *The Cerebellum as a Neuronal Machine* (Springer, New York, 1967). doi:10.1007/978-3-662-13147-3.
2. Ito, M. Movement and thought: identical control mechanisms by the cerebellum. *Trends in Neurosciences* **16**, 448–450 (1993).
3. Schmahmann, J. From movement to thought: Anatomic substrates of the cerebellar contribution to cognitive processing. *Human Brain Mapping* **4**, 174–198 (1996).
4. Habas, C. Functional Connectivity of the Cognitive Cerebellum. *Frontiers in Systems Neuroscience* **15**, 642225 (2021).
5. Palesi, F. *et al.* The Importance of Cerebellar Connectivity on Simulated Brain Dynamics. *Frontiers in Cellular Neuroscience* **14**, (2020).
6. Srivastava, P., Fotiadis, P., Parkes, L. & Bassett, D. S. The expanding horizons of network neuroscience: From description to prediction and control. *Neuroimage* **258**, 119250 (2022).
7. Koziol, L. F. *et al.* Consensus paper: The cerebellum's role in movement and cognition. *The Cerebellum* **13**, 151–177 (2014).
8. Mariën, P. *et al.* Consensus paper: Language and the cerebellum: An ongoing enigma. *Cerebellum* **13**, 386–410 (2014).
9. Caligiore, D. *et al.* Consensus Paper: Towards a Systems-Level View of Cerebellar Function: the Interplay Between Cerebellum, Basal Ganglia, and Cortex. *Cerebellum* **16**, 203–229 (2017).
10. Van Overwalle, F. *et al.* Consensus Paper: Cerebellum and Social Cognition. *The Cerebellum* **19**, 833–868 (2020).
11. Adamaszek, M. *et al.* Consensus Paper: Cerebellum and Emotion. *Cerebellum* **16**, 552–576 (2017).
12. Leto, K. *et al.* Consensus Paper: Cerebellar Development. *The Cerebellum 2015 15:6* **15**, 789–828 (2015).
13. Arleo, A. *et al.* Consensus Paper: Cerebellum and Ageing. *Cerebellum* (2023) doi:10.1007/s12311-023-01577-7.
14. Middleton, F. A. & Strick, P. L. Cerebellar projections to the prefrontal cortex of the primate. *Journal of Neuroscience* **21**, 700–712 (2001).
15. Kelly, R. M. & Strick, P. L. Cerebellar loops with motor cortex and prefrontal cortex of a nonhuman primate. *Journal of Neuroscience* **23**, 8432–8444 (2003).
16. Dum, R. P. & Strick, P. L. An unfolded map of the cerebellar dentate nucleus and its projections to the cerebral cortex. *Journal of Neurophysiology* **89**, 634–639 (2003).
17. Bostan, A. C. & Strick, P. L. The basal ganglia and the cerebellum: Nodes in an integrated network. *Nature Reviews Neuroscience* **19**, 338–350 (2018).
18. Ramnani, N. The primate cortico-cerebellar system: Anatomy and function. *Nature Reviews Neuroscience* **7**, 511–523 (2006).
19. Ramnani, N. *et al.* The evolution of prefrontal inputs to the cortico-pontine system: Diffusion imaging evidence from macaque monkeys and humans. *Cerebral Cortex* **16**, 811–818 (2006).
20. Palesi, F. *et al.* Contralateral cerebello-thalamo-cortical pathways with prominent involvement of associative areas in humans in vivo. *Brain Structure and Function* **220**, 3369–3384 (2015).
21. Palesi, F. *et al.* Contralateral cortico-ponto-cerebellar pathways reconstruction in humans in vivo: Implications for reciprocal cerebro-cerebellar structural connectivity in motor and non-motor areas. *Scientific Reports* **7**, 1–13 (2017).
22. Chauvel, M. Singularity of the white matter structural connectivity of the human brain compared to the chimpanzee brain. (Université Paris-Saclay, 2023).
23. Chauvel, M. *et al.* In vivo mapping of the deep and superficial white matter connectivity in the chimpanzee brain. *Neuroimage* **282**, 120362 (2023).

24. Magielse, N., Heuer, K., Toro, R., Schutter, D. J. L. G. & Valk, S. L. A Comparative Perspective on the Cerebello-Cerebral System and Its Link to Cognition. *Cerebellum* (2022) doi:10.1007/s12311-022-01495-0.
25. Magielse, N. *et al.* Phylogenetic comparative analysis of the cerebello-cerebral system in 34 species highlights primate-general expansion of cerebellar crura I-II. *Commun Biol* **6**, 1–17 (2023).
26. Glickstein, M. & Voogd, J. Cerebellum: Evolution and comparative anatomy. in *Encyclopedia of Neuroscience* (ed. Squire, L.) 743–756 (Academic Press Ltd., London, 2009).
27. Bolk, L. *Das Cerebellum Der Säugetiere: Eine Vergleichend Anatomische Untersuchung*. (Fischer, Jena, 1906).
28. Voogd, J. & Glickstein, M. The anatomy of the cerebellum. *Trends in Neurosciences* **21**, 370–375 (1998).
29. Larsell, O. & Jansen, J. *The Comparative Anatomy and Histology of the Cerebellum from Myxinoids through Birds. Annals of Internal Medicine* vol. 67 (University of Minnesota Press, Minneapolis, 1967).
30. Larsell, O. & Jansen, J. *The Comparative Anatomy and Histology of the Cerebellum: Vol. 2. From Monotremes through Apes.* vol. 2 (University of Minnesota Press, Minneapolis, 1970).
31. Nieuwenhuys, R. Comparative Anatomy of the Cerebellum. *Progress in Brain Research* **25**, 1–93 (1967).
32. Marr, D. *Vision: A Computational Investigation of Visual Representation in Man*. vol. 8 (Freeman and Company, WH, San Francisco, 1982).
33. Marr, D. A theory of cerebellar cortex. *The Journal of Physiology* **202**, 437–470 (1969).
34. Albus, J. S. A theory of cerebellar function. *Mathematical Biosciences* **10**, 25–61 (1971).
35. Ito, M. Neurophysiological aspects of the cerebellar motor control system. *International journal of neurology* **7**, 162–176 (1970).
36. Ito, M. & Kano, M. Long-lasting depression of parallel fiber-Purkinje cell transmission induced by conjunctive stimulation of parallel fibers and climbing fibers in the cerebellar cortex. *Neuroscience Letters* **33**, 253–258 (1982).
37. Ito, M., Sakurai, M. & Tongroach, P. Climbing fibre induced depression of both mossy fibre responsiveness and glutamate sensitivity of cerebellar Purkinje cells. *The Journal of Physiology* **324**, 113–134 (1982).
38. Schmahmann, J. Disorders of the cerebellum: Ataxia, dysmetria of thought, and the cerebellar cognitive affective syndrome. *Journal of Neuropsychiatry and Clinical Neurosciences* **16**, 367–378 (2004).
39. Diedrichsen, J., King, M., Hernandez-Castillo, C., Sereno, M. & Ivry, R. B. Universal Transform or Multiple Functionality? Understanding the Contribution of the Human Cerebellum across Task Domains. *Neuron* **102**, 918–928 (2019).
40. Orban de Xivry, J.-J. & Diedrichsen, J. Diversity of the nature of input and output signals in the cerebellum suggests a diversity of function. *Current Opinion in Behavioral Sciences* **57**, 101386 (2024).
41. Wang, D., Buckner, R. L. & Liu, H. Cerebellar asymmetry and its relation to cerebral asymmetry estimated by intrinsic functional connectivity. *Journal of Neurophysiology* **109**, 46–57 (2013).
42. Wolpert, D. M., Miall, R. C. & Kawato, M. Internal models in the cerebellum. *Trends in Cognitive Sciences* **2**, 338–347 (1998).
43. Ishikawa, T., Tomatsu, S., Izawa, J. & Kakei, S. The cerebro-cerebellum: Could it be loci of forward models? *Neuroscience Research* **104**, 72–79 (2016).
44. Tanaka, H., Ishikawa, T., Lee, J. & Kakei, S. The Cerebro-Cerebellum as a Locus of Forward Model: A Review. *Frontiers in Systems Neuroscience* **14**, 19 (2020).
45. Azevedo, F. A. C. *et al.* Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain. *Journal of Comparative Neurology* **513**, 532–541 (2009).
46. Tomasch, J. The numerical capacity of the human cortico-pontocerebellar system. *Brain Research* **13**, 476–484 (1969).

47. Hansel, C., Linden, D. J. & D'Angelo, E. Beyond parallel fiber LTD: the diversity of synaptic and non-synaptic plasticity in the cerebellum. *Nat Neurosci* **4**, 467–475 (2001).

48. Kawato, M., Ohmae, S., Hoang, H. & Sanger, T. 50 Years Since the Marr, Ito, and Albus Models of the Cerebellum. *Neuroscience* **462**, 151–174 (2020).

49. Bostan, A. C., Dum, R. P. & Strick, P. L. The basal ganglia communicate with the cerebellum. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 8452–8456 (2010).

50. Bostan, A. C., Dum, R. P. & Strick, P. L. Cerebellar networks with the cerebral cortex and basal ganglia. *Trends in Cognitive Sciences* **17**, 241–254 (2013).

51. Doya, K. What are the computations of the cerebellum, the basal ganglia and the cerebral cortex? *Neural Networks* **12**, 961–974 (1999).

52. Shine, J. M. & Shine, R. Delegation to automaticity: The driving force for cognitive evolution? *Frontiers in Neuroscience* **8**, 90 (2014).

53. Heuer, K., Traut, N., Sousa, A. A. de, Valk, S. & Toro, R. Diversity and evolution of cerebellar folding in mammals. 2022.12.30.522292 Preprint at https://doi.org/10.1101/2022.12.30.522292 (2022).

54. Diedrichsen, J., Balsters, J. H., Flavell, J., Cussans, E. & Ramnani, N. A probabilistic MR atlas of the human cerebellum. *NeuroImage* **46**, 39–46 (2009).

55. Marek, S. *et al.* Spatial and Temporal Organization of the Individual Human Cerebellum. *Neuron* **100**, 977-993.e7 (2018).

56. Nettekoven, C. *et al.* A hierarchical atlas of the human cerebellum for functional precision mapping. *Nat Commun* **15**, 8376 (2024).

57. Schlerf, J., Wiestler, T., Verstynen, T. & Diedrichsen, J. Big Challenges from the 'Little Brain'-Imaging the Cerebellum. in *Advanced Brain Neuroimaging Topics in Health and Disease* (eds. Papageorgiou, T. D., Christopoulos, G. I. & Smirnakis, S.) 199–223 (InTech, Rijeka, 2014).

58. Mathiesen, C., Caesar, K. & Lauritzen, M. Temporal coupling between neuronal activity and blood flow in rat cerebellar cortex as indicated by field potential analysis. *Journal of Physiology* **523**, 235–246 (2000).

59. Howarth, C., Peppiatt-Wildman, C. M. & Attwell, D. The energy use associated with neural computation in the cerebellum. *J Cereb Blood Flow Metab* **30**, 403–414 (2010).

60. King, M., Hernandez-Castillo, C. R., Poldrack, R. A., Ivry, R. B. & Diedrichsen, J. Functional boundaries in the human cerebellum revealed by a multi-domain task battery. *Nature Neuroscience* **22**, 1371–1378 (2019).

61. Zhi, D., King, M., Hernandez-Castillo, C. R. & Diedrichsen, J. Evaluating brain parcellations using the distance-controlled boundary coefficient. *Human Brain Mapping* **43**, 3706–3720 (2022).

62. Zhi, D. *et al.* A hierarchical Bayesian brain parcellation framework for fusion of functional imaging datasets. 2023.05.24.542121 Preprint at https://doi.org/10.1101/2023.05.24.542121 (2023).

63. Buckner, R. L., Krienen, F. M., Castellanos, A., Diaz, J. C. & Thomas Yeo, B. T. The organization of the human cerebellum estimated by intrinsic functional connectivity. *Journal of Neurophysiology* **106**, 2322–2345 (2011).

64. Xue, A. *et al.* The Detailed Organization of the Human Cerebellum Estimated by Intrinsic Functional Connectivity Within the Individual. *Journal of Neurophysiology* **125**, 358–384 (2020).

65. Guell, X., Gabrieli, J. D. E. & Schmahmann, J. D. Triple representation of language, working memory, social and emotion processing in the cerebellum: convergent evidence from task and seed-based resting-state fMRI analyses in a single large cohort. *NeuroImage* **172**, 437–449 (2018).

66. Eickhoff, S. B. *et al.* Coordinate-based activation likelihood estimation meta-analysis of neuroimaging data: a random-effects approach based on empirical estimates of spatial uncertainty. *Hum Brain Mapp* **30**, 2907–2926 (2009).

67. Eickhoff, S. B., Bzdok, D., Laird, A. R., Kurth, F. & Fox, P. T. Activation Likelihood Estimation meta-analysis revisited. *Neuroimage* **59**, 2349–2361 (2012).

68. Laird, A. *et al.* ALE meta-analysis: Controlling the false discovery rate and performing statistical contrasts. *Human Brain Mapping* **25**, 155–164 (2005).

69. Turkeltaub, P. E., Eden, G. F., Jones, K. M. & Zeffiro, T. A. Meta-analysis of the functional neuroanatomy of single-word reading: method and validation. *Neuroimage* **16**, 765–780 (2002).

70. Petacchi, A., Laird, A. R., Fox, P. T. & Bower, J. M. Cerebellum and auditory function: an ALE meta-analysis of functional neuroimaging studies. *Human brain mapping* **25**, (2005).

71. Baumann, O. & Mattingley, J. B. Scaling of neural responses to visual and auditory motion in the human cerebellum. *J Neurosci* **30**, 4489–4495 (2010).

72. Emch, M., von Bastian, C. C. & Koch, K. Neural Correlates of Verbal Working Memory: An fMRI Meta-Analysis. *Frontiers in Human Neuroscience* **13**, (2019).

73. Niendam, T. A. *et al.* Meta-analytic evidence for a superordinate cognitive control network subserving diverse executive functions. *Cogn Affect Behav Neurosci* **12**, 241–268 (2012).

74. Radua, J., Pozo, N. O. D., Gómez, J., Guillen-Grima, F. & Ortuño, F. Meta-analysis of functional neuroimaging studies indicates that an increase of cognitive difficulty during executive tasks engages brain regions associated with time perception. *Neuropsychologia* **58**, 14–22 (2014).

75. Bernard, J. & Seidler, R. Cerebellar contributions to visuomotor adaptation and motor sequence learning: an ALE meta-analysis. *Frontiers in Human Neuroscience* **7**, (2013).

76. Stoodley, C. J. & Schmahmann, J. D. Functional topography in the human cerebellum: A meta-analysis of neuroimaging studies. *NeuroImage* **44**, 489–501 (2009).

77. E, K., Chen, S. A., Ho, M. R. & Desmond, J. E. A meta-analysis of cerebellar contributions to higher cognition from PET and fMRI studies. *Hum Brain Mapp* **35**, 593–615 (2012).

78. Van Overwalle, F., Ma, Q. & Heleven, E. The posterior crus II cerebellum is specialized for social mentalizing and emotional self-experiences: a meta-analysis. *Soc Cogn Affect Neurosci* **15**, 905–928 (2020).

79. Van Overwalle, F., Baetens, K., Mariën, P. & Vandekerckhove, M. Social cognition and the cerebellum: A meta-analysis of over 350 fMRI studies. *NeuroImage* **86**, 554–572 (2014).

80. Pierce, J. E., Thomasson, M., Voruz, P., Selosse, G. & Péron, J. Explicit and Implicit Emotion Processing in the Cerebellum: A Meta-analysis and Systematic Review. *Cerebellum* (2022) doi:10.1007/s12311-022-01459-4.

81. Kruithof, E. S., Klaus, J. & Schutter, D. J. L. G. The human cerebellum in reward anticipation and outcome processing: An activation likelihood estimation meta-analysis. *Neuroscience & Biobehavioral Reviews* **149**, 105171 (2023).

82. Baumann, O. & Mattingley, J. B. Functional topography of primary emotion processing in the human cerebellum. *NeuroImage* **61**, 805–811 (2012).

83. Balsters, J. H., Laird, A. R., Fox, P. T. & Eickhoff, S. B. Bridging the gap between functional and anatomical features of cortico-cerebellar circuits using meta-analytic connectivity modeling. *Human Brain Mapping* **35**, 3152–3169 (2014).

84. Balsters, J. H., Whelan, C. D., Robertson, I. H. & Ramnani, N. Cerebellum and Cognition: Evidence for the Encoding of Higher Order Rules. *Cerebral Cortex* **23**, 1433–1443 (2013).

85. Riedel, M. C. *et al.* Meta-analytic connectivity and behavioral parcellation of the human cerebellum. *NeuroImage* **117**, 327–342 (2015).

86. Van Overwalle, F. *et al.* A Functional Atlas of the Cerebellum Based on NeuroSynth Task Coordinates. *Cerebellum* (2023) doi:10.1007/s12311-023-01596-4.

87. Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C. & Wager, T. D. Large-scale automated synthesis of human functional neuroimaging data. *Nat Methods* **8**, 665–670 (2011).

88. Laird, A. *et al.* ALE meta-analysis workflows via the BrainMap database: progress towards a probabilistic functional brain atlas. *Frontiers in Neuroinformatics* **3**, (2009).

89. Desmond, J. E. & Fiez, J. A. Neuroimaging studies of the cerebellum: language, learning and memory. *Trends in Cognitive Sciences* **2**, 355–362 (1998).

90. Fox, P. T. *et al.* Brainmap taxonomy of experimental design: Description and evaluation. *Human Brain Mapping* **25**, 185–198 (2005).

91. Fonov, V., Evans, A., McKinstry, R., Almli, C. & Collins, D. Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. *NeuroImage* **47**, S102 (2009).

92. Larsell, O. & Jansen, J. *The Comparative Anatomy and Histology of the Cerebellum: The Human Cerebellum, Cerebellar Connections, and Cerebellar Cortex*. *Neurology* vol. 23 (University of Minnesota Press, Minneapolis, 1973).

93. Schmahmann, J., Doyon, J., Toga, A., Petrides, M. & Evans, A. MRI Atlas of the Human cerebellum. (2000).

94. Guell, X., Schmahmann, J. D., Gabrieli, J. D. E. & Ghosh, S. S. Functional gradients of the cerebellum. *eLife* **7**, e36652 (2018).

95. Ji, J. L. *et al.* Mapping the human brain's cortical-subcortical functional network organization. *NeuroImage* **185**, 35–57 (2019).

96. Eickhoff, S. B. *et al.* Behavior, sensitivity, and power of activation likelihood estimation characterized by massive empirical simulation. *Neuroimage* **137**, 70–85 (2016).

97. Langner, R., Rottschy, C., Laird, A. R., Fox, P. T. & Eickhoff, S. B. Meta-analytic connectivity modeling revisited: controlling for activation base rates. *Neuroimage* **99**, 559–570 (2014).

98. Tian, Y., Margulies, D. S., Breakspear, M. & Zalesky, A. Topographic organization of the human subcortex unveiled with functional connectivity gradients. *Nature Neuroscience 2020 23:11* **23**, 1421–1432 (2020).

99. García-Cabezas, M. Á., Zikopoulos, B. & Barbas, H. The Structural Model: a theory linking connections, plasticity, pathology, development and evolution of the cerebral cortex. *Brain Structure and Function* **224**, 985–1008 (2019).

100. Diedrichsen, J. A spatially unbiased atlas template of the human cerebellum. *NeuroImage* **33**, 127–138 (2006).

101. Markello, R. D. *et al.* neuromaps: structural and functional interpretation of brain maps. *Nat Methods* **19**, 1472–1479 (2022).

102. Vos de Wael, R. *et al.* BrainSpace: a toolbox for the analysis of macroscale gradients in neuroimaging and connectomics datasets. *Communications Biology 2020 3:1* **3**, 1–10 (2020).

103. Adrian, F. D. Localization in the Cerebrum and Cerebellum. *British Medical Journal* **2**, 137–140 (1944).

104. Manoli, A., Overwalle, F. V., Wiesmann, C. G. & Valk, S. L. Functional recruitment and connectivity of the cerebellum supports the emergence of Theory of Mind in early childhood. 2024.04.02.586955 Preprint at https://doi.org/10.1101/2024.04.02.586955 (2024).

105. Balsters, J. H. *et al.* Evolution of the cerebellar cortex: The selective expansion of prefrontal-projecting cerebellar lobules. *NeuroImage* **49**, 2045–2052 (2010).

106. Strick, P. L., Dum, R. P. & Fiez, J. A. Cerebellum and Nonmotor Function. *Annual Review of Neuroscience* **32**, 413–434 (2009).

107. Bush, E. C. & Allman, J. M. The scaling of frontal cortex in primates and carnivores. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 3962–3966 (2004).

108. Habas, C. *et al.* Distinct cerebellar contributions to intrinsic connectivity networks. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **29**, 8586–8594 (2009).

109. Watson, T. C. *et al.* Anatomical and physiological foundations of cerebello-hippocampal interaction. *eLife* **8**, e41896 (2019).

110. Jung, S. J. *et al.* Novel Cerebello-Amygdala Connections Provide Missing Link Between Cerebellum and Limbic System. *Frontiers in Systems Neuroscience* **16**, (2022).

111. Terburg, D., van Honk, J. & Schutter, D. J. L. G. Doubling down on dual systems: A cerebellum–amygdala route towards action- and outcome-based social and affective behavior. *Cortex* **173**, 175–186 (2024).

112. Zeidler, Z., Hoffmann, K. & Krook-Magnuson, E. HippoBellum: Acute Cerebellar Modulation Alters Hippocampal Dynamics and Function. *J. Neurosci.* **40**, 6910–6926 (2020).

113. Katsumi, Y. *et al.* Correspondence of functional connectivity gradients across human isocortex, cerebellum, and hippocampus. *Commun Biol* **6**, 1–13 (2023).

114. Liu, X. *et al.* A multifaceted gradient in human cerebellum of structural and functional development. *Nature Neuroscience* **25**, 1129–1133 (2022).

115. Margulies, D. S. *et al.* Situating the default-mode network along a principal gradient of macroscale cortical organization. *Proceedings of the National Academy of Sciences* **113**, 12574–12579 (2016).

116. Paquola, C. *et al.* Microstructural and functional gradients are increasingly dissociated in transmodal cortices. *PLoS Biology* **17**, (2019).

117. D'Mello, A. M., Gabrieli, J. D. E. & Nee, D. E. Evidence for Hierarchical Cognitive Control in the Human Cerebellum. *Current Biology* (2020) doi:10.1016/j.cub.2020.03.028.

118. King, M., Zhen, Z., Ivry, R. B. & Weiner, K. S. Transcriptomic Gradients Of The Human Cerebellum. SSRN Scholarly Paper at https://doi.org/10.2139/ssrn.3797269 (2021).

119. Wang, Y. *et al.* Spatio-molecular profiles shape the human cerebellar hierarchy along the sensorimotor-association axis. *Cell Reports* **43**, 113770 (2024).

120. Straub, I. *et al.* Gradients in the mammalian cerebellar cortex enable Fourier-like transformation and improve storing capacity. *eLife* **9**, e51771 (2020).

121. King, M., Shahshahani, L., Ivry, R. B. & Diedrichsen, J. A task-general connectivity model reveals variation in convergence of cortical inputs to functional regions of the cerebellum. *eLife* **12**, e81511 (2023).

122. Apps, R. *et al.* Cerebellar Modules and Their Role as Operational Cerebellar Processing Units. *The Cerebellum* **17**, 654–682 (2018).

123. Cerminara, N. L. & Apps, R. Behavioural significance of cerebellar modules. *Cerebellum* **10**, 484–494 (2011).

124. Cerminara, N. L., Lang, E. J., Sillitoe, R. V. & Apps, R. Redefining the cerebellar cortex as an assembly of non-uniform Purkinje cell microcircuits. *Nature Reviews Neuroscience 2015 16:2* **16**, 79–93 (2015).

125. Müller, V. I. *et al.* Ten simple rules for neuroimaging meta-analysis. *Neuroscience & Biobehavioral Reviews* **84**, 151–161 (2018).

126. Tahmasian, M. *et al.* Practical recommendations to conduct a neuroimaging meta-analysis for neuropsychiatric disorders. *Human Brain Mapping* **40**, 5142–5154 (2019).

127. Priovoulos, N. & Bazin, P.-L. Methods for cerebellar imaging analysis. *Current Opinion in Behavioral Sciences* **54**, 101328 (2023).

128. Vaidya, M. V. *et al.* Improved Detection of fMRI Activation in the Cerebellum at 7T with Dielectric Pads Extending the Imaging Region of a Commercial Head Coil. *Journal of magnetic resonance imaging* **48**, 431 (2018).

129. Priovoulos, N., Andersen, M., Dumoulin, S. O., Boer, V. O. & van der Zwaag, W. High-Resolution Motion-corrected 7.0-T MRI to Derive Morphologic Measures from the Human Cerebellum in Vivo. *Radiology* 220989 (2023) doi:10.1148/radiol.220989.

130. Vanasse, T. J. *et al.* BrainMap VBM: An environment for structural meta-analysis. *Human Brain Mapping* **39**, 3308–3325 (2018).

131. Turkeltaub, P. E. *et al.* Minimizing within-experiment and within-group effects in Activation Likelihood Estimation meta-analyses. *Hum Brain Mapp* **33**, 1–13 (2012).

132. Salo, T. *et al.* NiMARE: Neuroimaging Meta-Analysis Research Environment. *Aperture Neuro* **3**, 1–32 (2023).

133. Salo, T. *et al.* neurostuff/NiMARE: 0.2.0rc3. Zenodo https://doi.org/10.5281/zenodo.8359277 (2023).

134. Viladomat, J., Mazumder, R., McInturff, A., McCauley, D. J. & Hastie, T. Assessing the significance of global and local correlations under spatial autocorrelation: A nonparametric approach. *Biometrics* **70**, 409–418 (2014).

135. Burt, J. B., Helmer, M., Shinn, M., Anticevic, A. & Murray, J. D. Generative modeling of brain maps with spatial autocorrelation. *NeuroImage* **220**, 117038 (2020).

136. Alexander-Bloch, A. F. *et al.* On testing for spatial correspondence between maps of human brain structure and function. *NeuroImage* **178**, 540–551 (2018).

137. Guell, X. *et al.* Littlebrain: A gradient-based tool for the topographical interpretation of cerebellar neuroimaging findings. *PLoS ONE* **14**, e0210028 (2019).

138. Saberi, A. *et al.* The regional variation of laminar thickness in the human isocortex is related to cortical hierarchy and interregional connectivity. *PLOS Biology* **21**, e3002365 (2023).