

## Reviewed Preprint

v1 • June 10, 2026

Not revised

## ✉ For correspondence:

satra@mit.edu

**Competing interests:** No competing interests declared**Funding:** See [page 39](#)**Reviewing editor:** Yanchao Bi, Peking University, China

© 2026, Mentch et al. This article is distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use and redistribution provided that the original author and source are credited.

# Pregistered movie-fMRI analyses reveal altered visual feature encoding in autism in pSTS

Jeff Mentch<sup>1,2</sup>, Yibei Chen<sup>2</sup>, Tamara Vanderwal<sup>3,4</sup>, Satrajit S Ghosh<sup>1,2,5</sup> ✉

<sup>1</sup>Program in Speech and Hearing Bioscience and Technology, Harvard University, Cambridge, United States • <sup>2</sup>McGovern Institute for Brain Research, Massachusetts Institute of Technology, Cambridge, United States • <sup>3</sup>Department of Psychiatry, University of British Columbia, Vancouver, Canada; • <sup>4</sup>BC Children's Hospital Research Institute, Vancouver, Canada • <sup>5</sup>Department of Otolaryngology, Harvard Medical School, Boston, United States

## eLife Assessment

This **valuable** study uses naturalistic movie-viewing fMRI and stacked encoding models to investigate sensory feature representations in autistic and non-autistic youth, showing a relative shift toward low-level visual representations in higher-order social cortical regions in autism. The evidence is **solid** overall, supported by preregistration, a relatively large open dataset, and sophisticated encoding-model analyses, although several methodological and interpretive issues require further clarification and validation. The work will interest researchers in developmental cognitive neuroscience and naturalistic neuroimaging.

<https://doi.org/10.7554/eLife.111008.1.sa3>

## Abstract

Sensory-perceptual differences are widely reported in autism, yet their underlying mechanisms remain unclear. We tested preregistered hypotheses using stacked encoding models applied to naturalistic movie-viewing fMRI from children and adolescents with and without an autism diagnosis from the Healthy Brain Network. We mapped cortical responsiveness to low- and high-level auditory and visual feature spaces. Contrary to enhanced perceptual functioning predictions, autism was not associated with increased low-level encoding in primary sensory cortices. Instead, autistic children and adolescents had reduced high-level visual representations and a relative shift toward low-level over high-level feature encoding in integration and social brain regions including the pSTS and adjacent face/social areas. In pSTS, this high-low weighting tracked Social Responsiveness Scale (SRS) scores. By contrast, audio-visual modality preference and sensory dominance were broadly conserved across groups. Developmentally, encoding exhibited strong, lateralized, modality-congruent age effects. Together, these findings favor weak central coherence accounts over early sensory enhancement, constrain mechanisms to altered visual feature weighting within social/multisensory networks, and demonstrate the value of naturalistic stimuli and encoding models for characterizing sensory-perceptual neurodevelopmental differences.

## Introduction

Sensory processing differences are a common and clinically meaningful feature of autism with a profound impact on daily life (Ben-Sasson et al., 2007 [↗](#); Chamak et al., 2008 [↗](#)). As many as 90% of autistic individuals exhibit altered sensory function (Leekam et al., 2007 [↗](#); Tavassoli et al., 2014 [↗](#)) with differences reported across all senses (Robertson and Baron-Cohen, 2017 [↗](#); Dakin and Frith, 2005 [↗](#); Nieto Del Rincón, 2008 [↗](#); Bennetto et al., 2007 [↗](#); Balasco et al., 2020 [↗](#); DuBois et al., 2016 [↗](#)), including features such as hyper- and hypo-sensitivities, enhanced perceptual abilities (Lefebvre et al., 2023 [↗](#); Mottron et al., 2006 [↗](#); O'riordan, 2004 [↗](#); Mottron and Burack,

2001), altered susceptibility to perceptual illusions (Happé, 1996; Feldman et al., 2022), and differences in multisensory integration (MSI) (Stevenson et al., 2014b,a; Feldman et al., 2018). These traits emerge early, track with social and cognitive outcomes, and vary across individuals and modalities (Uljarević et al., 2017; Williams et al., 2023; Georgiades et al., 2017; Chen et al., 2022; Simonoff et al., 2020; Smith et al., 2012). Despite their prevalence and impact, the neurobiological substrates of these altered sensory functions remain relatively uncharacterized (Marco et al., 2011; Robertson and Baron-Cohen, 2017).

## Theories of sensory differences

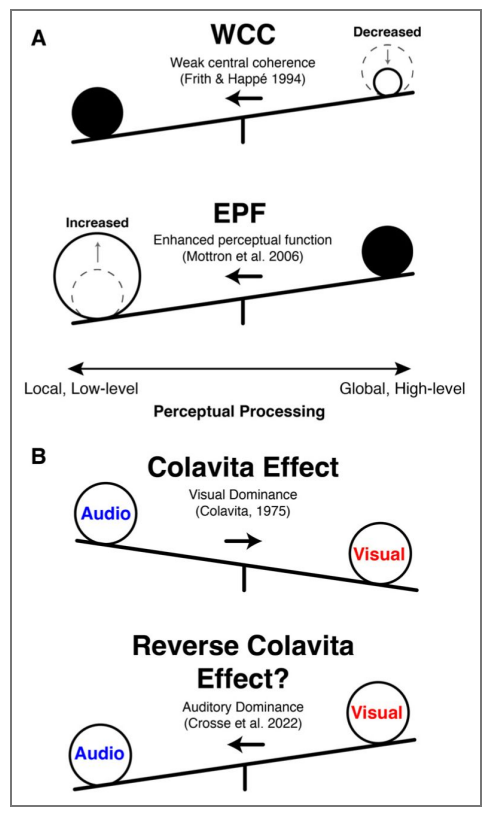
A number of different theories exist surrounding autistic sensory differences. Enhanced perceptual functioning (EPF) emphasizes heightened low-level processing and improved discrimination (Mottron and Burack, 2001; Mottron et al., 2006), while weak central coherence (WCC) highlights a local-over-global perceptual preference (Frith, 1989; Happé et al., 2006) (see Figure 1 A for an illustrated comparison). Related Bayesian accounts propose altered use or updating of perceptual priors or an atypical balance between statistical likelihood and priors (Pellicano and Burr, 2012; Schneebeli et al., 2022; Brock, 2012) and further, predictive-coding, a process-level instantiation of these Bayesian ideas, implicates imbalances between prediction errors and top-down predictions (Mumford, 1992; Rao and Ballard, 1999; Friston, 2005; Lawson et al., 2014; Van de Cruys et al., 2014; Lawson et al., 2017). More mechanistically, neurobiological accounts emphasize excitation–inhibition (E/I) imbalances and neural noise, with support from EEG, genetics, and animal studies (Rubenstein and Merzenich, 2003; Dinstein et al., 2012; Puts et al., 2017; Wei et al., 2021; Hollestein et al., 2023; Han et al., 2012; Robertson et al., 2016). An integrative view links E/I alterations to disrupted divisive normalization central to causal inference and predictive coding (Noel and Angelaki, 2023). In recent years, “Sensory-first” accounts of autism argue that perceptual differences precede and scaffold downstream social–communicative phenotypes (Levit-Binnun et al., 2013; Cascio et al., 2016; Robertson and Baron-Cohen, 2017; Falck-Ytter and Bussu, 2023; Russo et al., 2025; Takesian and Hensch, 2013).

Overall, these varied accounts yield some testable predictions for naturalistic representations, among them: relatively amplified low-level and reduced high-level encoding and altered local–global and unimodal–multisensory balance;

Another axis of sensory organization relevant to autism is the balance between auditory and visual channels during naturalistic perception. Classic multisensory work shows that audiovisual (AV) integration engages regions including the superior temporal sulcus (STS), which can be supraadditive for congruent AV events (Calvert et al., 2000, 2001), with multisensory influences present even in regions often treated as unisensory (Ghazanfar and Schroeder, 2006; Stein and Stanford, 2008). Developmentally, AV integration is protracted, approaching adult-like levels around late childhood, and may follow altered trajectories in autism (Neil et al., 2006; Burr and Gori, 2012; Russo et al., 2025). Behaviorally, autism has been associated with differences in multisensory integration and temporal binding (e.g., widened temporal binding windows and reduced MSI enhancements), with downstream implications for domains including speech processing (Stevenson et al., 2014b,a; Feldman et al., 2018; Wang et al., 2024). Classic work on sensory dominance shows a robust visual-dominance in adults (the Colavita effect) that develops from earlier auditory predominance in childhood (Colavita, 1974; Hirst et al., 2018; Ross et al., 2021). In autism, findings include reduced or reversed visual dominance, including a reverse Colavita effect and a shift toward auditory bias (Ross et al., 2021; Moro et al., 2012; O’Connor and Hermelin, 1965), hinting at potentially altered audio vs. visual sensory feature weightings in autism (see Figure 1 B).

## Developmental maturation of sensory and association systems

Beyond autism, it is useful to note several robust developmental trends. Primary sensory cortices mature relatively early, whereas higher-order association regions (e.g., pSTS, TPJ, FFA) show protracted structural–functional specialization into adolescence (Thompson and Nelson, 2001);



**Figure 1.**

(A) A visual depiction contrasting how perceptual processing is thought to be different in autism across the two prominent theories of WCC and EPF. (B) An illustration comparing visual sensory dominance (the Colavita effect) with decreased visual dominance or auditory sensory dominance (a reverse Colavita effect), which has been observed in autism.

Sydnor et al., 2021). Multisensory processing also sharpens with age: audiovisual temporal binding windows narrow, and cross-modal influences in STS become more reliable and congruent with task demands (Neil et al., 2006; Burr and Gori, 2012). These maturational trends imply that group differences observed in association cortices and multisensory networks must be interpreted against a baseline of ongoing developmental tuning.

## Neuroimaging evidence for altered sensory processing in autism

Within autism, neuroimaging findings have consistently implicated primary sensory brain regions (e.g. motion-related differences in visual cortex (Robertson et al., 2014), enlarged population receptive fields (Schwarzkopf et al., 2014), and heightened connectivity within primary visual and auditory cortices (Martínez et al., 2020; Cerliani et al., 2015)). Increased trial-to-trial variability of evoked responses has also been observed across early visual, auditory, and somatosensory cortices (Dinstein et al., 2012). Atypical early auditory discrimination has been further measured through mismatch negativity (MMN), with MMN amplitudes directly linked to sensory over-responsivity and autistic traits (Cary et al., 2024). During audiovisual integration tasks, autistic participants were found to exhibit auditory-led phase resets in visual cortex, opposite the visual-led synchronization typical of neurotypical observers, consistent with altered precision and reliability weighting of sensory cues and tying primary sensory processing alterations to audiovisual perception (Ronconi et al., 2023).

Alongside these early sensory processing differences, a growing body of research highlights significant involvement of higher-order cortical regions. These include but are not limited to the temporoparietal junction (TPJ; found to have decreased cortical thickness (You et al., 2024) and altered connectivity to the cerebellum (Igelström et al., 2017)), the posterior superior temporal sulcus (pSTS; found to have reduced activation during biological motion perception (Koldewyn et al., 2011; Federici et al., 2020) and altered intrinsic and dynamic functional connectivity (Alaerts et al., 2015, 2014; Shih et al., 2011; Guo et al., 2019)), and the fusiform face area (FFA; found to have altered volume and hemispheric asymmetry (Herbert et al., 2002; Floris et al., 2025) along with reduced activation in response to faces (Pierce et al., 2001; Schultz et al., 2003; Hadjikhani et al., 2007), though some later studies found no difference for familiar faces (Pierce et al., 2004; Hadjikhani et al., 2007)).

Beyond specific low- and high-level regions with altered connectivity, there have been widespread reports of more generally altered connectivity across the brain, in the form of under-connectivity (Damarla et al., 2010; Abrams et al., 2013; Alaerts et al., 2014), over-connectivity (Rudie and Dapretto, 2013; Keown et al., 2013), and differing local and global connectivity (Courchesne and Pierce, 2005). These patterns are thought to be dynamic and changing with age (Lawrence et al., 2019). Some findings point to an anterior to posterior gradient across the brain of under- to over-connectivity in autism, perhaps explaining enhanced perceptual performance with increased local connectivity in early sensory areas and disrupted social behavior with decreased local connectivity in frontal areas (Keown et al., 2013; Supekar et al., 2013; Di Martino et al., 2014; Rudie and Dapretto, 2013). While most studies assess activation or connectivity, fewer examine what or how information is represented during rich, dynamic perception and a clear gap remains.

## Naturalistic stimuli and autism

Naturalistic stimuli (such as movies and narratives) in neuroimaging studies can improve ecological validity and compliance while also implicitly including rich sensory, semantic, and social features (Sonkusare et al., 2019; Nastase et al., 2020; Vanderwal et al., 2019; Redcay and Moraczewski, 2020; Haxby et al., 2014). In autism, movie paradigms have revealed reduced ISC and ISFC and idiosyncratic responses (foundational and follow-up work by Hasson et al. 2009; Salmi et al. 2013; Byrge et al. 2015), links to comprehension and symptomatology (e.g., default-mode/control networks; Bolton and colleagues (2018; 2020)), region- or event-specific effects (TPJ, theory-of-mind; Pantelis et al. 2015; Lyons et al. 2020), and cross-national replication of ISFC reductions (Lin et al., 2025). Deficits in audiovisual synchrony relevant to

speech have also been reported (Wang et al., 2024). Most work to date leverages metrics of brain synchrony, while feature-resolved accounts of what and where information may be differentially represented in these naturalistic contexts remains relatively unexplored.

## Encoding models

Encoding models enable the direct mapping of stimulus features to brain responses, allowing us to quantify the representation of information across the brain (Naselaris et al., 2011; Diedrichsen and Kriegeskorte, 2017). Encoding models are readily adaptable to high-dimensional, naturalistic features (Samara et al., 2025; Khosla et al., 2021; Huth et al., 2016; Haxby et al., 2011) and some specific implementations, such as stacked encoding models (Lin et al., 2024) and banded regression (Dupré la Tour et al., 2022) enable joint tests across feature spaces. In stacked encoding models, instead of relying solely on the predictive performance of separate feature models, the secondary “stacking” stage yields weights that index the relative contribution of each feature space to the joint prediction, providing, for example, a direct measure of high- vs. low-level or audio vs. visual perceptual preferences when these feature classes compete to explain the same responses. Despite their successful applications in vision and language research, encoding model approaches are relatively underutilized in autism research relative to ISC based methods, particularly in studies using naturalistic designs (Dupré la Tour et al., 2025; Kriegeskorte and Douglas, 2019).

## This study: dataset, features, analyses, preregistration, hypotheses

Here, naturalistic movie-viewing fMRI data from children and adolescents with and without autism is analyzed to test a set of preregistered hypotheses (accessible at <http://osf.io/h92gr> and <http://osf.io/47kj6>; summarized in Table 1). Feature sets are constructed spanning low-level (e.g., brightness, motion and loudness) through high-level (faces, bodies, and audio categories such as speech and music), separately for audio and visual streams and feed these into whole-brain grayordinate-level encoding models that quantify the explained variance for each feature class. Primary hypotheses: (i) Relative to non-autistic peers, autistic participants will show greater encoding of *low-level* perceptual features in primary sensory and other perceptual regions (H1.1), with a shift in stacked-weight preference toward lowover high-level features (H1.2); (ii) autistic auditory and visual cortices will be more unimodal (H2.1), preferentially encoding modality-congruent information and exhibiting reduced visual dominance (H2.2). Secondary hypotheses: Encoding-model metrics will covary with clinical phenotype including autism severity and sensory symptoms (H1.3, H2.3).

	Hypothesis	Feature sets / metrics
<b>Low- vs. high-level encoding</b>		
1.1	Primary perceptual cortices will encode more information about low-level perceptual features in an autistic relative to a non-autistic group.	Low-level audio and visual features; $R^2$ , $R_u^2$
1.2	Perceptual regions of the brain will be more tuned towards low- as opposed to high-level sensory information in an autistic relative to a non-autistic group.	Low- vs. high- level audio and visual feature classes; $(W_{HA} - W_{LA})$ , $(W_{HV} - W_{LV})$
1.3	Encoding model-derived differences will be related to autism severity and sensory symptoms.	SRS/SSS associations with high- and low-level, audio and visual $R^2$ , $R_u^2$ , $(W_H - W_L)$
<b>Audio vs. visual encoding</b>		
2.1	Auditory and visual regions of the brain will be more <i>unimodal</i> in autistic compared to non-autistic people and encode more congruent unimodal information relative to incongruent cross-modal information.	Audio-only vs. visual-only contributions; audio and visual $R^2$ , $R_u^2$
2.2	Autistic individuals will have reduced visual dominance compared to non-autistic individuals.	Visual vs. audio stacked weights; $(W_V - W_A)$
2.3	Audiovisual encoding model-derived differences in autistic participants will be significantly correlated with autism severity.	SRS/SSS associations with audio and visual $R^2$ , $R_u^2$ and $(W_V - W_A)$

**Table 1.** Table of preregistered hypotheses (accessible at [osf.io/h92gr](http://osf.io/h92gr) and [osf.io/47kj6](http://osf.io/47kj6)).

## Key findings

During naturalistic viewing fMRI, autistic children and adolescents have a relative shift toward low-level over high-level feature encoding in integration and social brain regions (including the pSTS), with a broadly conserved audio–visual feature balance but robust age-related effects. These findings inform sensory theories, argue against large, pervasive shifts in modality dominance, and demonstrate the utility of naturalistic encoding model methods for the investigation of sensory-perceptual neurodevelopmental differences.

## Results

### Conventions

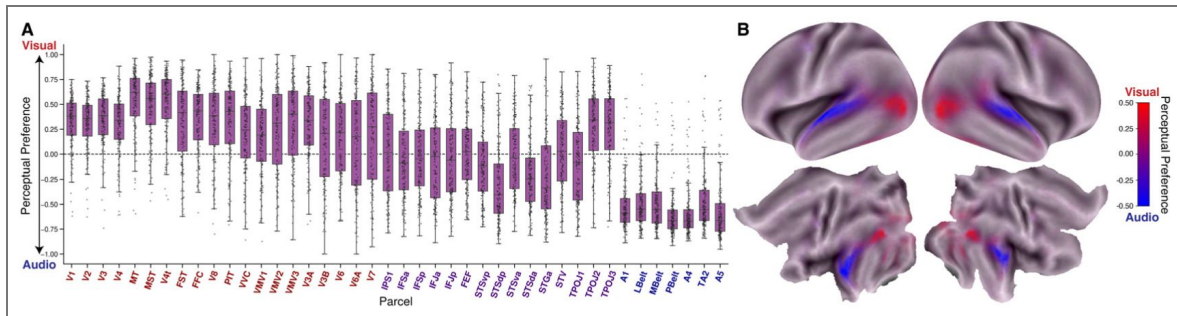
We report explained variance ( $R^2$ ), unique explained variance ( $R_u^2$ , from variance partitioning), and stacked encoding model weights ( $W$ ) for audio and visual models. Perceptual preference indices are defined as  $W_V - W_A$  (visual vs. audio modality preference) and  $W_H - W_L$  (high- vs. low-level feature preference), with positive values indicating visual (or high-level) preference and negative values indicating audio (or low-level) preference. Unless otherwise noted, all p-values are FDR-corrected across parcels, and results are shown across three framewise displacement (FD) thresholds of 40%, 60%, and 80%.

### Encoding models recover known sensory hierarchies

Aggregating across all (both autistic and non-autistic) participants, audiovisual stacked encoding models produced patterns of performance and feature representation in audio and visual cortical regions that are in keeping with known sensory hierarchies. This is evident by examining the perceptual preference index (the difference in visual and auditory model stacked weights ( $W_V - W_A$ )), across both perceptual ROIs and the cortical surface more broadly (Figure 2). Mean  $R^2$  peaked along the superior temporal sulcus (STS), and split-half, Spearman–Brown–corrected noise ceilings (an upper bound on explainable variance given response reliability/measurement noise) were highest in parcels along the STS and in parietal/occipital cortices (Figure 2—figure Supplement 1). Looking at model performance by sensory modality, auditory cortices showed greater audio  $R^2$  and increased audio model weights, whereas visual cortices showed the converse pattern for visual performance and weights (Figure 2—figure Supplement 2). Comparing stacked low- and high-level models within audio and visual models further reproduced canonical hierarchical gradients. For vision, low-level features dominated early visual cortex (e.g., V1/V2), with relatively stronger high-level contributions in downstream ventral regions (e.g., FFC, Figure 2—figure Supplement 3 A). For the audition, high-level model  $R^2$  and weights increased from A1 into association areas (A4/A5, Figure 2—figure Supplement 3 B). Model weights followed this same pattern along with the high vs. low preference index ( $W_H - W_L$ ), which was negative in early sensory cortex and approached or exceeded zero in higher-order regions (Figure 2—figure Supplement 3 C–F, see Figure 2—figure Supplement 4 for a whole-brain mean visualization). These results establish that the feature spaces and model architecture employed here are capable of recovering known sensory hierarchies.

### H1.1: No evidence for increased low-level encoding in primary sensory cortices in autism

The preregistered hypothesis that primary perceptual cortices contain more information about low-level perceptual features in autism was not supported. For the low-level auditory model, no significant group differences in  $R^2$  or  $R_u^2$  were observed in primary auditory regions (all FDR  $q > 0.05$ , Figure 3, Figure 3—figure Supplement 1). Outside auditory ROIs, some uncorrected effects appeared for  $R_u^2$  (V4t, MST) and  $R^2$  (V4t, V7), but none survived FDR correction. Likewise, no significant group-level differences in low-level visual performance were observed in primary visual regions (all FDR  $q > 0.05$ , Figure 3, Figure 3—figure Supplement 2). Trends in STSvp,



**Figure 2.**

(A) Visual vs audio perceptual preference ( $W_V - W_A$ ) for all participants (ASD and nonASD) across visual (red), audiovisual (purple), and auditory (blue) ROIs. Perceptual preference is calculated by taking the difference of visual and audio stacked model weights. Perceptual preference is in the range -1 to 1 as stacked encoding model weights range from 0 to 1. (B) Whole-brain plot of mean visual vs audio perceptual preference ( $W_V - W_A$ ) across all participants (ASD and nonASD). This metric is the same as in A, but now plotted at each grayordinate and colored on a scale of visual (red) to audio (blue) with purple indicating regions with a relative balance between visual and audio feature weights.

STSdp, and STGa (uncorrected  $p < .05$ ) did not survive FDR correction for  $R^2$ , whereas STSvp showed a significant difference for  $R_u^2$  (40% FD; replicated at 80% FD; Figure 3). Overall, early sensory-cortex encoding increases for low-level features in autism were not observed.

### H1.2: In autism, pSTS is weighted towards low-level visual features

Although audio models had no significant difference in low- vs. high-level preference, stacked visual models revealed a group difference in high- vs. low-level preference localized to the pSTS. Visual  $W_H - W_L$  was significantly lower in the autistic group in STSvp across all three FD thresholds and in STSdp at the strictest 40% threshold (Figure 3, Figure 4).

### H1.3: pSTS encoding differences are related to autism severity as measured by the SRS

Across visual and auditory models and the SRS and SSS measures, visual high- vs. low-level preference was significantly related to SRS after FDR correction in STSdp at the 40% threshold where lower  $W_H - W_L$  (increased low-level preference) was associated with higher SRS score (coefficients summarized in Figure 3).

### Exploratory: High-level visual representations are reduced in face and social brain regions in autism

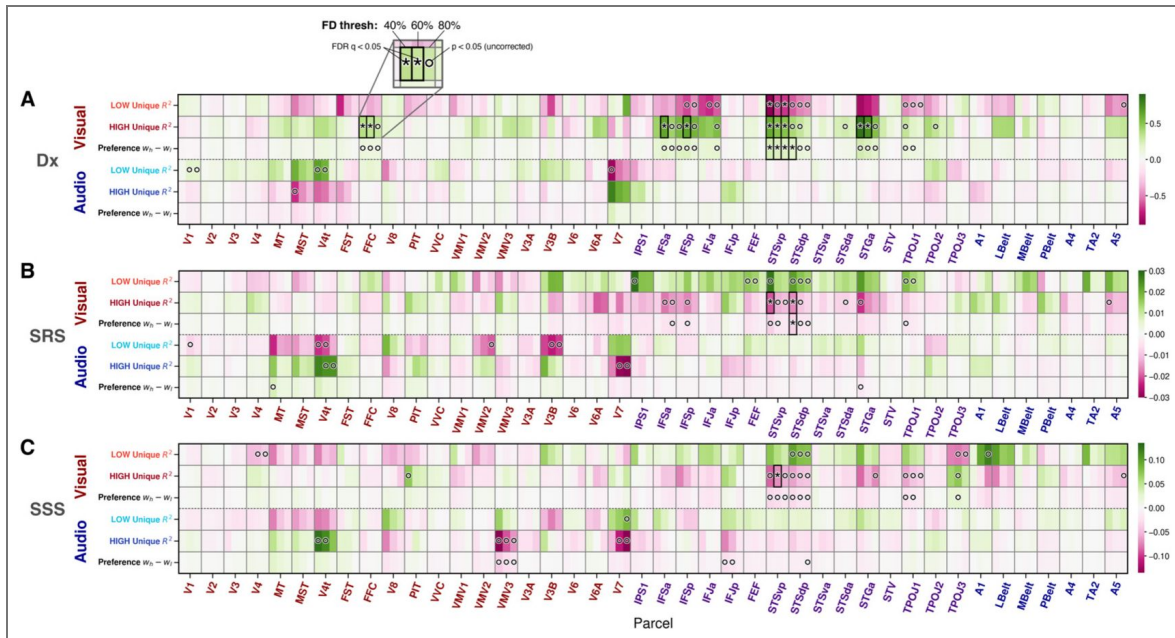
Although the preregistered hypotheses focused on low-level feature representation and the balance of low- and high-level feature weights, the stacked models also provide high-level feature representation metrics. High-level audio  $R^2$  and  $R_u^2$  showed no group differences (Figure 3—figure Supplement 3). By contrast, high-level visual features showed reliable group effects in face- and social-perception related regions including FFC, STS, STG, and inferior frontal cortex. The effects were largely consistent across both  $R^2$  and  $R_u^2$  (Figure 3—figure Supplement 4), indicating robust differences in high-level visual information that are not explained by variance shared with low-level features.

### Exploratory: Whole-brain parcel-wise visual encoding differences and ASD±ADHD stratification

Whole-brain low- and high-level stacked encoding model analyses identified additional parcels beyond the a priori ROIs with significant visual encoding differences. At the 40% FD threshold, low-level visual  $R_u^2$  was greater in the autism group in left 8BL (lateral BA8, posterior superior frontal gyrus; DLPFC;  $q < 0.005$ ) and right SFL ( $q < 0.05$ ); across thresholds, left 8BL also showed higher high-level preference in the non-autistic group (Table 2; cortical maps in Figure 5). Other parcels with significant visual  $R^2/R_u^2$  differences at one or more thresholds included right TE1p, right PF, left IFSa, left STSvp, and striatal regions CAU-DA-lh and CAU-body-lh (Table 2).

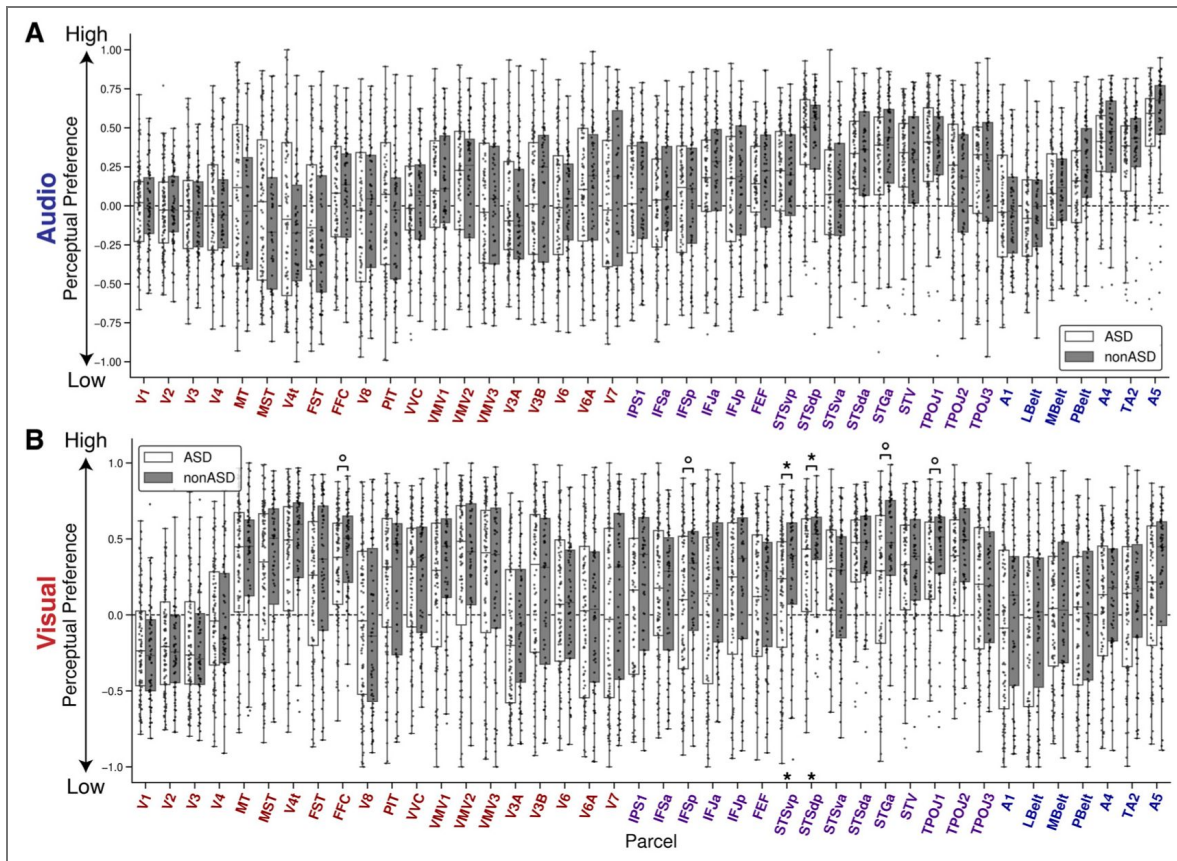
	Right_SFL	Right_TE1p	Right_PF	Left_8BL	Left_6r	Left_IFSa	Left_STSvp	CAU-DA-lh	CAU-body-lh
VL $R^2$	-	-	-	-	-	-	-	*, -	-
VL $R_u^2$	*, -	*, *	-	***, *	-	-	-	-	-
VH $R^2$	*, -	*, *	-	*, *	-	-	-	-	-
VH $R_u^2$	-	*, *	*, -	*, *	-	*, -	*, -	-	*, -
Pref. ( $W_{HV} - W_{LV}$ )	-	-	*, -	*, *	-	*, -	*, -	-	*, -
AH $R_u^2$	-	-	-	-	**, -	-	-	-	-

**Table 2. Table of post-hoc whole-brain significance results.** Only significant regions and features (after FDR correction) are included in the table. In each column, results from FD thresholds of 40%, 60%, and 80% are separated by commas. VL=low-level visual, VH=high-level visual, AH=high-level audio,  $W_{VL}$ =low-level visual model weight,  $W_{VH}$ =high-level visual model weight,  $R^2$ =explained variance,  $R_u^2$ =unique explained variance, -= $q > 0.05$ , \*= $q < 0.05$ , \*\*= $q < 0.005$



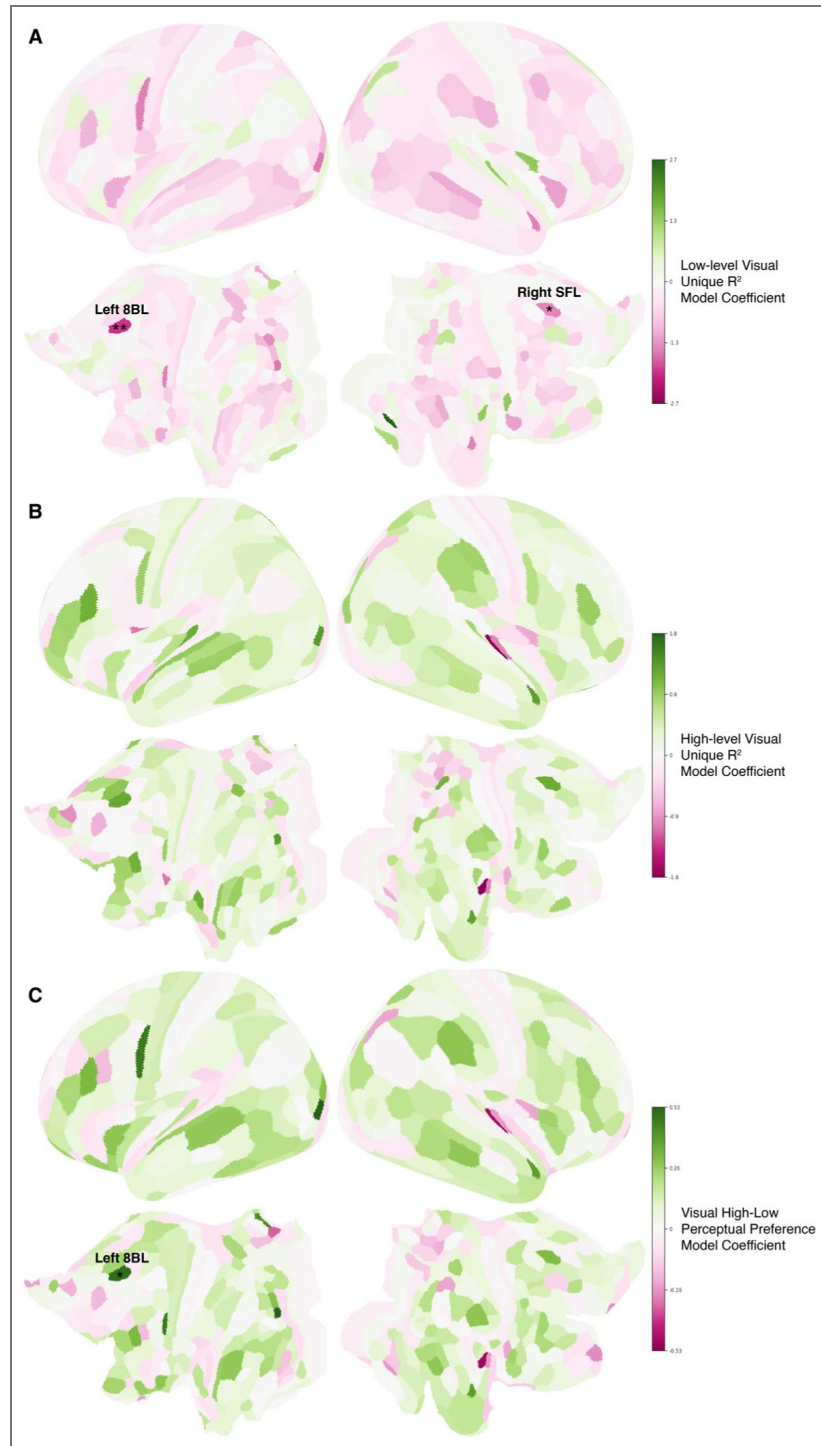
**Figure 3.** Heatmaps of fixed-effect coefficients for (A) diagnostic group (Dx: nonASD vs. ASD), (B) SRS and (C) sensory subset score (SSS), across 40%, 60% and 80% FDR thresholds (left to right within each cortical parcel column).

Within each of the three horizontal panels, rows denote encoding model-derived metrics (low-level  $R_U^2$ , high-level  $R_U^2$ , and their preference index ( $W_H - W_L$ ) and columns denote Glasser ROIs ordered from early visual areas through association cortex and then auditory areas. Color indicates the magnitude and sign of the coefficient (purple=negative effect with ASD>nonASD; green=positive effect with ASD<nonASD). Asterisks (\*) mark FDR-corrected significance at  $q < .05$ ; open circles (o) mark uncorrected  $p < .05$ . The visual modality encoding models are coded with red text and the auditory with blue. Low-level visual  $R_U^2$  was significantly different across diagnosis in STSvp and low-level audio  $R_U^2$  was not significant anywhere (A; contradicting H1.1). Visual though not audio  $W_H - W_L$  was significantly different in STSvp and STSdp (A; supporting H1.2). Visual but not audio  $W_H - W_L$  was significantly related to SRS in STSdp at the 40% threshold (B; H1.3) but not SSS (C). Generally, results were often consistent across the three FDR thresholds.



**Figure 4.**

(A) Box plot of audio stacked encoding model perceptual preference across ASD (white) and nonASD (gray) groups for all perceptual ROIs. (B) Corresponding box plot of the visual stacked encoding models. Results correspond to the 40% FD threshold. Boxes annotated with an asterisk indicate a significant group difference (FDR  $q < 0.05$ ) while a circle indicates an initially significant difference between groups that did not survive FDR correction. Boxes show the quartiles of the dataset and whiskers show the distribution with the exception of outliers. Each dot is the mean perceptual preference from a single subject from all statistically significant grayordinates within each ROI.



**Figure 5.** Cortical surface maps showing visual encoding model-derived fixed-effect coefficients for diagnostic group comparisons (ASD vs. nonASD) at the 40% FD threshold.

Colors represent coefficient magnitude and direction (pink: negative effect, ASD > nonASD; green: positive effect, ASD < nonASD). Asterisks denote significance after FDR correction (\*: FDR  $q < 0.05$ , \*\*: FDR  $q < 0.005$ ).

Given the high ADHD comorbidity in the sample (and in autism generally), the autism group was further subdivided into ASD+ADHD and ASD-ADHD subgroups. Across perceptual ROIs, significant differences in visual encoding metrics were observed between nonASD and ASD-ADHD participants, but not between nonASD and ASD+ADHD (note the nonASD control group has no participants with a diagnosis of ADHD), and there were no differences between ASD subgroups (Table 3). The ASD-ADHD vs. nonASD results largely mirrored the broader autistic group at the ROI level as well as at the whole-brain level (e.g. a significant difference in left 8BL). A robust auditory model finding also emerged at the whole-brain level in the left Middle Insular area (MI) for low-level audio  $R_u^2$  at 40% FD in ASD+ADHD vs. nonASD.

	IFSp	STSvp	STSdp
<b>VL <math>R^2</math> (<math>NT, ADHD^-</math>)</b>	-	-, -, *	-
<b>VL <math>R_u^2</math> (<math>NT, ADHD^-</math>)</b>	-, *, -	*, * **	-, *, -
<b>VH <math>R^2</math> (<math>NT, ADHD^-</math>)</b>	*, -, -	*, * *	*, -, -
<b>VH <math>R_u^2</math> (<math>NT, ADHD^-</math>)</b>	-	-, -, *	-
<b>Pref. (<math>W_{VL} - W_{VH}</math>) (<math>NT, ADHD^-</math>)</b>	-	-, *, **	-

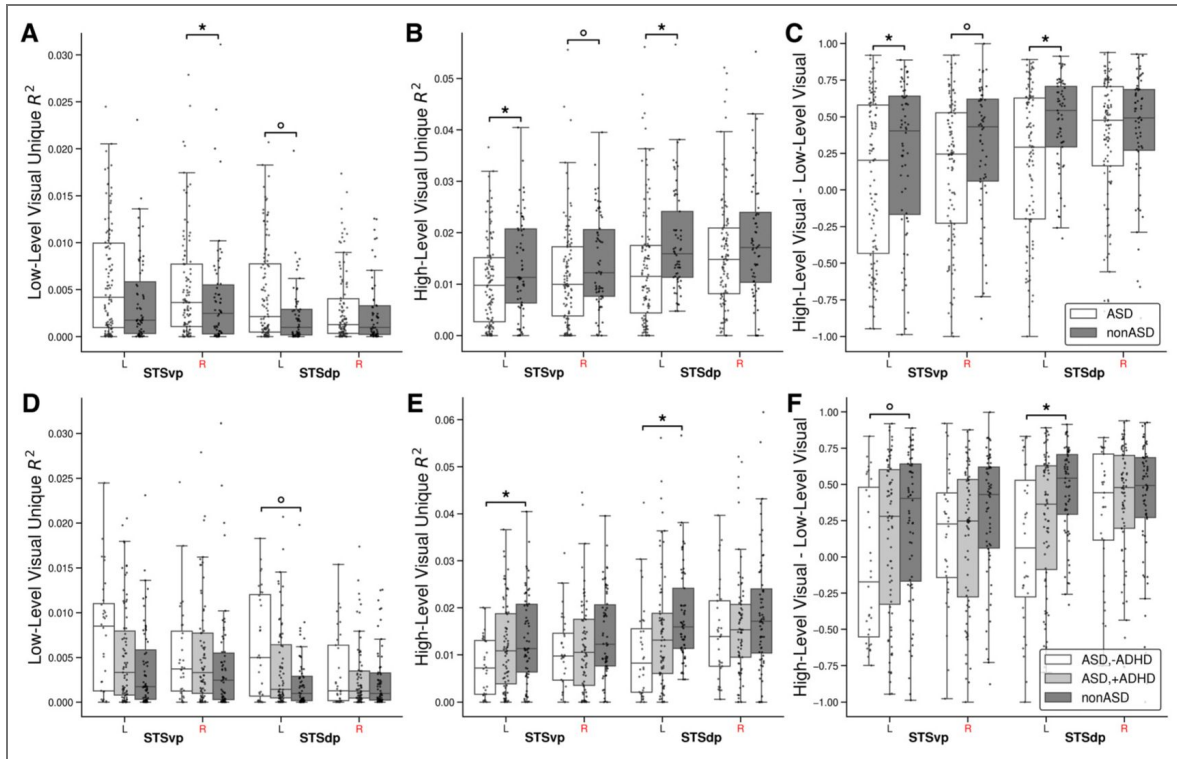
**Table 3. Perceptual regions of interest ADHD subgroup significance summary.** All effects listed here are between the nonASD and ADHD+ASD subgroups, as there were no significant differences between nonASD and ADHD-ASD and ADHD+ASD and ADHD-ASD. Only the three regions with significant effects after FDR correction are included. In each column, the results from FD thresholds of 40%, 60% and 80% are separated by columns. VL=low-level visual, VH=high-level visual

### Exploratory: pSTS left-hemisphere predominance with ASD±ADHD corroboration

Given converging effects in the pSTS, hemispheric laterality in that region was examined in a posthoc analysis. For low-level visual features, group differences in  $R_u^2$  were evident in right STSvp but not left (Figure 6 A). In contrast, for high-level visual  $R_u^2$  and  $W_H - W_L$ , group differences were stronger and more consistent in the left hemisphere (STSvp and STSdp; Figure 6 B,C). Right STSvp trended in the same direction for both  $R_u^2$  and  $W_H - W_L$  but did not survive FDR at the 40% FD threshold (uncorrected  $p < .05$ ). Taken together, this pattern suggests a hemispheric division of effects. ADHD-stratified analyses corroborated this laterality. Comparing ASD-ADHD with nonASD reproduced the pSTS effects with the same left-hemisphere predominance across metrics (Figure 6 D-F).

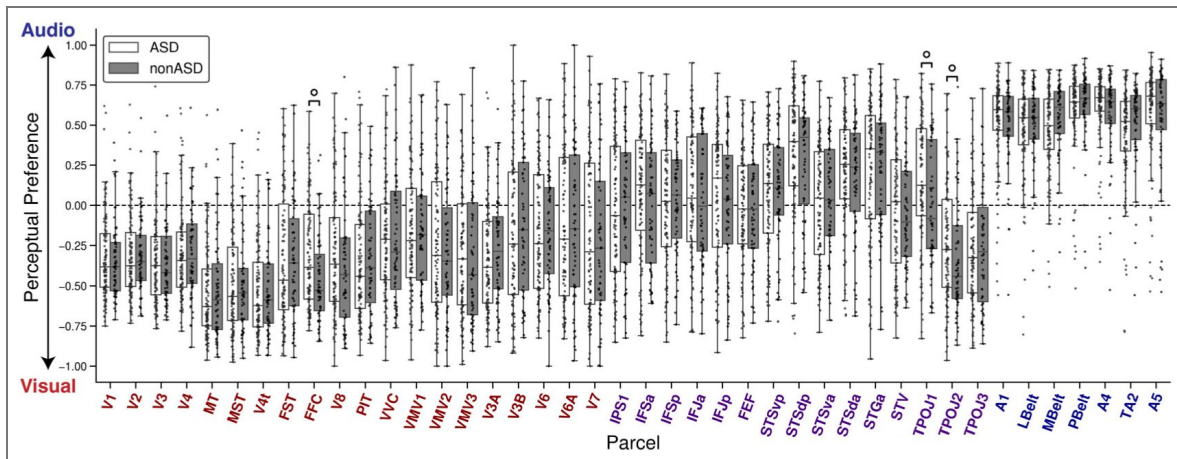
### H2: Audio-visual feature representation and modality preference is conserved across groups and not related to autism severity

There were no significant ASD vs. nonASD differences in visual vs. audio preference ( $W_V - W_A$ , Figure 7) or audio and visual  $R^2$  (Figure 7—figure Supplement 1) or  $R^2$  (Figure 7—figure Supplement 2) within their corresponding cortical regions, with the singular exception of TPOJ2 at the 60% FD threshold (greater visual  $R^2$  in the nonASD relative to ASD group, Figure 7—figure Supplement 3). Overall we observed minimal evidence of reduced visual dominance (or increased auditory dominance) related to audio and visual feature representation in autism during naturalistic viewing. Beyond diagnostic group differences, audio and visual model  $R^2$ ,  $R_u^2$  and  $W_V - W_A$  were also not significantly related to SRS or SSS (Figure 7—figure Supplement 3 B,C).



**Figure 6. Visual encoding model metrics across left and right pSTS subregions.**

(A) Low-level visual unique explained variance ( $R_U^2$ ) for ASD (white) and nonASD (gray) groups in left and right STSvp and STSdp regions. (B) High-level visual  $R_U^2$  across the same groups and regions. (C) High- vs. low-level perceptual preference (derived from the stacked visual encoding models), where positive values indicate a high-level preference, negative values indicate a low-level preference, and values near zero indicate no preference. (D) Low-level visual  $R_U^2$  for ASD-ADHD (white), ASD+ADHD (light gray), and nonASD (gray) groups across left and right STSvp and STSdp regions. (E) High-level visual  $R_U^2$  across groups in the same regions. (F) Stacked encoding model visual perceptual preference metrics (high- vs. low-level); positive values indicate a preference for high-level features, negative values indicate a low-level preference, and values near zero indicate no strong perceptual preference. All results are from the 40% FD threshold. Boxes annotated with an asterisk (\*) indicate significant group differences (FDR-corrected  $p < 0.05$ ); circles indicate nominal significance prior to FDR correction (uncorrected  $p < 0.05$ ). Boxes represent quartiles; whiskers indicate data range excluding outliers. Individual dots represent mean metrics per subject, averaged across statistically significant grayordinates within each left or right ROI. Pairwise significance tests were conducted between all groups; annotations indicate significant differences, all of which occurred between the ASD-ADHD and nonASD subgroups.



**Figure 7. Audio vs. visual modality preference.**

Higher values indicate an audio preference and lower values indicate a visual preference. These results correspond with the 40% FD threshold. Boxes annotated with circles indicate a difference between groups that did not survive FDR correction (uncorrected  $p < 0.05$  but  $FDR q > 0.05$ ). Boxes show the quartiles of the dataset and whiskers show the distribution with the exception of outliers. Each dot is the mean value from a single subject from all statistically significant grayordinates within each ROI.

### Exploratory: Age exerts prominent effects on modality preference

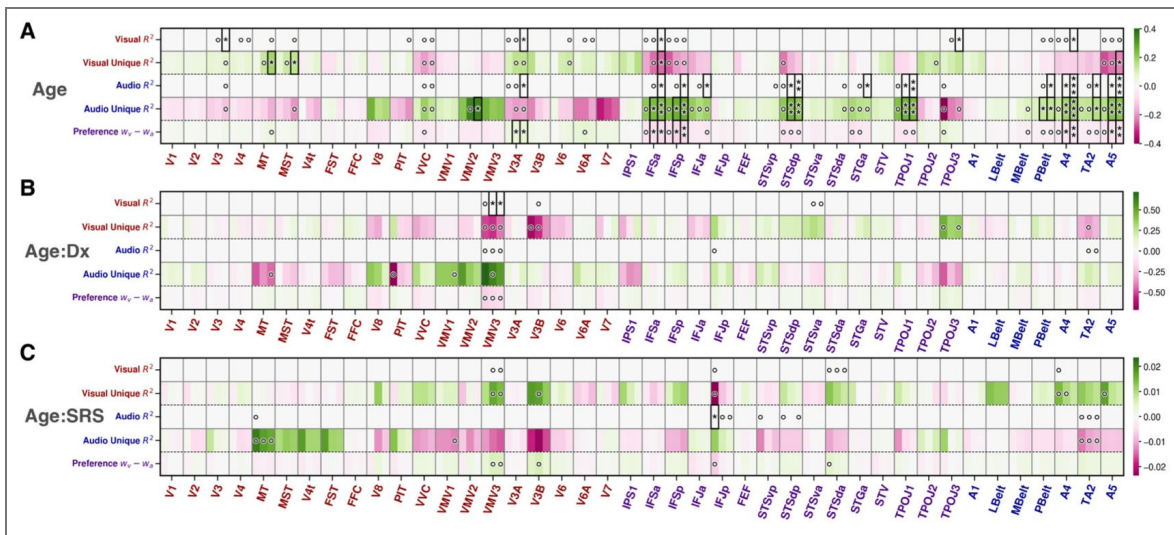
Focusing on age without regard to diagnostic group, robust relationships emerged across perceptual ROIs (Figure 8). Visual  $R^2$  and  $R_u^2$  increased with age in V3, MT, MST, V3A, IFSa, TPOJ3, and A4; auditory  $R^2$  and  $R_u^2$  increased with age in VMV2, V3A, IFSa, IFSp, IFJa, STSdp, STGa, TPOJ1, PBelt, A4, TA2, and A5. Modality preference ( $W_V - W_A$ ) shifted with age in V3A, IFSa, IFSp, A4, and A5, such that the visual region V3A had an increase in visual preference and the other auditory and multisensory regions had an increased auditory preference. Notably, these age effects were prominent at the 60% and 80% FD thresholds but not at the 40% threshold, plausibly because this stricter criteria disproportionately excluded more younger participants due to their increased motion.

Whole-brain analyses (where all cortical and subcortical parcels were considered separately across the left and right hemispheres) confirmed many ROI effects and further revealed some hemispheric specificities (Table 4): for example, auditory  $R^2$  related to age in A5 bilaterally, whereas V3A/MT effects were right-lateralized and A4/STSdp/TPOJ1 were left-lateralized. Beyond the predefined perceptual ROIs, additional parcels with significant age relationships included bilateral LO3, right TGv, left 9a, and left RI, adjacent to primary auditory cortex (Glasser et al., 2016). Several language-network parcels (pSTS/55b/PSL/SFL/area 45) also showed age effects for auditory  $R^2$  and  $R^2$ , predominantly left-lateralized (Glasser et al., 2016). In all cases, the directionality matched modality (auditory metrics increased with age in auditory regions; visual metrics increased with age in visual regions), and modality preference became more congruent with regional specialization (See Figure 8—figure Supplement 1 and Figure 8—figure Supplement 2 for examples from perceptual and whole-brain regions respectively).

	Visual $R^2$	Visual unique $R^2$	Audio $R^2$	Audio unique $R^2$	Pref. ( $W_{\text{Visual}} - W_{\text{Audio}}$ )
Right V3A	-	-	-	- , - *	-
Right MT	-	- , - *	-	-	-
Right IFSp	-	-	- , - *	- , - *	-
Right OP4	-	-	- , - *	- , - *	-
Right A5	-	-	- , - *	- , - *	- , - *
Right LO3	-	-	-	* , - -	* , - -
Right VVC	-	- , * , -	-	-	-
Right TGv	-	-	-	- , * , -	-
Left 55b	-	-	- , - *	- , - *	-
Left PSL	-	-	- , - **	- , * , *	-
Left SFL	-	-	- , - *	- , * , *	-
Left 45	-	-	- , - *	-	-
Left 9a	-	-	-	* , - , -	-
Left RI	-	-	- * , **	- , * , *	-
Left A5	-	-	- , - *	- , - *	-
Left STSda	-	-	- , - *	- , - *	-
Left STSdp	-	-	- , - *	- , * , *	-
Left TPOJ1	-	-	- , - *	- , - *	-
Left DVT	* , - , -	-	-	* , - , -	* , - , -
Left LO3	-	-	-	* , - , -	-
Left A4	-	-	- , - **	- , * , **	- , - *

**Table 4. Significance of Age effects on visual, audio, and audio–visual encoding model metrics across all left and right hemisphere, cortical and subcortical parcels at the whole-brain level.**

Post-hoc whole-brain ADHD subgroup comparisons revealed one region with consistent auditory  $R^2$  differences that survived correction across all 410 parcels: left PoI2 (insula), with ASD–ADHD vs. noNASD differences across all three FD thresholds (FDR  $q = 0.03, 0.003, 0.008$  for 40%, 60%, and 80%, respectively). No other pairwise subgroup contrasts (ASD–ADHD vs. ASD+ADHD; ASD+ADHD vs. nonASD) yielded auditory effects that survived FDR correction.



**Figure 8.** Heatmaps of fixed-effect coefficients from exploratory parcel-wise linear mixed models examining (A) main effects of age, (B) age-by-diagnosis (ASD vs. nonASD), and (C) age-by-SRS interactions.

Within each horizontal panel, rows represent encoding model-derived metrics (visual  $R^2$  and  $R_u^2$ , audio  $R^2$  and  $R_u^2$ , and their perceptual preference index  $W_V - W_A$ ), while columns represent Glasser MMP parcels ordered roughly from early visual through higher-order association to auditory areas. Each parcel is further subdivided by FD thresholds (40%, 60%, 80%, left to right). Color indicates magnitude and direction of coefficients (green: positive effects; pink: negative effects). In panel A, green indicates encoding metrics increase with age, pink indicates metrics decrease with age. In panel B (Age  $\times$  Diagnosis), green indicates the age-related slope is more positive in nonASD (greater increases with age), whereas pink indicates a more positive age-related slope in ASD. Asterisks mark FDR-corrected significance (\* =  $q < 0.05$ , \*\* =  $q < 0.005$ , \*\*\* =  $q < 0.0005$ ); open circles mark uncorrected significance ( $p < 0.05$ ). Visual metrics are labeled in red; auditory metrics in blue.

## Exploratory: Age-by-diagnosis and Age-by-SRS interactions are sparse and localized

At the ROI level, only VMV3 showed a significant Age×Diagnosis interaction for visual  $R^2$  where in the autistic group, visual  $R^2$  was relatively stable across age, whereas in the non-autistic group it decreased with age (Figure 8—figure Supplement 3 [↗](#)). Whole-brain analyses identified two additional Age×Diagnosis interactions for visual  $R_u^2$ : right area 52 and left p32pr, where visual  $R_u^2$  increased with age in autism but decreased with age in non-autistic participants. For Age×SRS, only audio  $R^2$  in IFJp showed a significant interaction at the ROI level (autism: slight increase with age; nonautism: decrease), and at the whole-brain level, interactions appeared for visual  $R^2$  and  $R_u^2$  in right 7PL (the autistic group measure slightly decreasing with age while the non-autistic group measure increases with age).

### Results Summary

Overall, this study finds (i) no evidence for increased low-level encoding specific to primary sensory cortices in autism; (ii) reduced high-level visual feature representations in social and face brain regions (STSVp, STGa, FFC) in autism; (iii) a relative preference towards low-level visual features in the pSTS in autism, strongest in the left hemisphere and linked to SRS score; and (iv) conserved audio–visual modality dominance and feature encoding in autism. Further, results show that feature encoding becomes increasingly more modality-congruent with age, with lateralized parcelspecific trajectories, though interactions of age with diagnosis or SRS score are sparse. Post-hoc whole-brain and ADHD-stratified analyses further identify additional brain regions with significant encoding differences, including left 8BL and left PoI2.

### Discussion

We asked where and how sensory features are differentially represented in the brain during naturalistic viewing in autism. First, we found no evidence for stronger low-level encoding in primary sensory cortices in autism; instead, differences localized to higher-order association regions, most notably the pSTS and adjacent frontal parcels. Second, in the autistic group, pSTS showed a preference for lower-level visual information over high-level visual encoding, with perceptual preference significantly related to SRS scores. Third, audio–visual modality balance was broadly conserved in autism, with significant effects predominantly observed developmentally in relation to age rather than phenotypically across diagnostic groups.

### Implications for sensory theories

The observed pattern of downstream, visual-specific alterations, with reduced high-level representations in social and integration hubs, fits the theory of weak central coherence (WCC) more readily than enhanced perceptual functioning (EPF), which predicts increased encoding in early sensory cortex. At a minimum, high-level global integration differences appear to outweigh any enhancements of early visual coding in this dataset and experimental context. A hybrid account remains plausible, where both theories are true or where subtle early differences cascade to association regions, for example, but our results do not directly support this.

### pSTS, lateral pathway, and social perception

The emergence of the pSTS as a locus of autistic group differences in our results is mechanistically informative. The pSTS is known to support perceptual functions including audiovisual speech (Okada et al., 2013 [↗](#)), biological motion (Sokolov et al., 2012 [↗](#)), and higher-order social perception (Lahnakoski et al., 2012 [↗](#)), with changes observed across development (Xiao et al., 2016 [↗](#)). Here, autistic participants showed decreased high-level visual encoding and a relative preference toward low-level visual features, with a left-hemisphere predominance for the high-level reduction. This lateralized pattern sits within the context of the proposed third, lateral visual pathway thought to be specialized for social perception, and raises the possibility that altered visual components of socially relevant signals might underlie reported downstream effects, such

as reduced audiovisual speech perception, observed behaviorally in autism. One key detail is that the SRS association was specific to the perceptual preference measure ( $W_H - W_L$ ) rather than raw performance, pointing to an altered weighting of visual evidence rather than a uniform gain/loss. Another key observation is that high-level visual  $R^2$  itself was reduced in social brain regions (FFC/STS/STG), consistent with a selective reduction in category-level visual information and not a broader effect that might be attributable to something like increased noise.

## Audio–visual dominance: largely conserved

Contrary to predictions related to behavioral “reverse Colavita” reports, we did not observe group differences in audio–visual modality preference in sensory cortices. Naturalistic, passive viewing and ceiling/floor constraints in unimodal cortices may attenuate sensitivity to cross-modal shifts behaviorally measured under different contexts; moreover, the sensory dominance transitions reported in psychophysics may precede our sampled age range or require task demands that are absent here. An interpretation of these findings is that shifts in sensory modality dominance, at least at the levels measurable by fMRI-based encoding models, are not a hallmark of autistic perception during naturalistic movie viewing.

## Developmental effects dominate audio and visual encoding

Age explained more variance than the diagnostic group when considering the stacked audio-visual encoding models. Specifically, encoding performance increased with age within corresponding modalities (visual model in visual cortex; audio model in auditory cortex, multisensory, and language areas), and modality preference also became more congruent with regional specialization. Effects were often lateralized (e.g., right dorsal-stream visual parcels; left language network), consistent with known specializations and a general network-specific reweighting. These results demonstrate the ability of encoding models to pick up broad developmental effects while ruling out these same factors in relation to diagnosis and phenotype. Notably, most age-related effects appeared only at looser FD thresholds, and stricter FD thresholds disproportionately excluded younger participants. This highlights the importance of performing analyses with an awareness and transparency around factors such as noise and data quality, as this can be a critical factor that affects downstream results and inferences.

## ADHD comorbidity

Post-hoc stratification results suggest that visual encoding differences are driven by autistic participants without ADHD; the ASD+ADHD subgroup showed little divergence from the non-autistic group, with a lone insular auditory effect under the strictest motion threshold. This pattern motivates explicit modeling of ADHD comorbidity in future work and raises the possibility of potentially opposing sensory feature encoding in ASD vs. ADHD. This finding is particularly interesting, given that functional connectivity has been shown to predict SRS across both ADHD and ASD (Lake et al., 2019 [↗](#)).

## Sex-based considerations

Autism shows sex-linked differences in prevalence, diagnostic presentation, and neurobiology, with growing evidence for partly distinct neural phenotypes in females and males (Napolitano et al., 2022 [↗](#)). Although sex was included as a covariate, the present study was neither designed nor powered to adjudicate sex-specific encoding effects. Consequently, the pSTS findings and high- vs. low-level visual reweighting could be disproportionately driven by males. Future work should pre-register sex-stratified analyses (and Sex × Diagnosis and Sex × Age interactions), recruit adequately powered female cohorts, and incorporate pubertal status as a moderator. Such designs will clarify whether the observed reductions in high-level visual encoding within social–perceptual networks reflect a male-typical autistic phenotype, a shared mechanism across sexes, or qualitatively different pathways in autistic females.

## Limitations and Future Work

- Stimulus duration. The combined movie runs were comparatively brief, which limited encoding performance, constrained the feasible feature space (due to overfitting when increasing the number of features), and limited the ability to perform noise-ceiling and within-subject reliability estimation. We therefore used compact, theoretically robust features with crossvalidation and permutation testing. Future work using longer acquisitions (and/or repetitions) would permit richer models to probe for finer-grained effects.
- MSI-related inferences are limited by modality and design. With only mixed audiovisual runs, no unimodal (A-only/V-only) acquisitions, and no behavioral MSI measures (e.g., temporal binding window or Colavita task), direct tests of multisensory integration were not possible. Moreover, many audiovisual interactions unfold at millisecond scales beyond the resolution of the BOLD signal. We therefore used modality-resolved encoding and variance partitioning to separate audio and visual contributions within fMRI's constraints. Future work should incorporate unimodal runs and MSI tasks, and pair fMRI with EEG/MEG to capture fast crossmodal dynamics.
- Sensory phenotyping. No dedicated sensory-processing scales (e.g., SSP/SPM) were available to map our results directly to sensory phenotypes. Using the scales available in the dataset with sensory related questions (SRS-2 and SCQ) we aggregated and used a Sensory Subset Score (SSS), though it was not psychometrically validated. Future work should pair encoding model results with dedicated sensory-processing scales or validate the SSS.
- Sample composition and generalizability. The dataset included only verbally fluent participants and the autism cohort had a male-skewed sex distribution and high ADHD comorbidity, limiting generalizability, particularly excluding minimally and non-verbal autistic individuals and autistic females. We included age, sex, socioeconomic status, and site as confounds in all analyses, and in exploratory analyses, stratified by ADHD diagnosis, but future work should balance sex, ADHD, and verbal ability and model comorbidity a priori.
- Preprocessing/parcellation choices and lack of functional localizers. We used MMP parcels without full MSMAll alignment and without subject-specific functional localizers (e.g., a face localizer for FFA) as they were not available. This may have introduced some degree of registration imprecision and ROI mis-localization. We did perform MSMSulc cortical surface registration, allowing for improved surface alignment. Future work could apply MSMAll alignment, functional localizers, or subject-specific parcellations to improve anatomical specificity.
- Eye gaze and attention. Autistic differences in fixation patterns during naturalistic viewing are well documented and can alter effective visual input (Klin et al., 2002 [↗](#)). As eye-tracking data was not directly collected during fMRI acquisition and regressors were not gaze-weighted, we cannot exclude eye-movement confounding or that our observed differences may be due to autistic participants watching the movie differently. This warrants future work incorporating eye-tracking and gaze-weighted features.
- Network-level dynamics and directionality. Our analyses were regional and did not assess how feature representations are coordinated across networks or the direction of information flow. Within these constraints we limited inference to parcel-wise metrics. Future work should relate encoding to static and time-varying functional connectivity and apply effectiveconnectivity approaches to test directionality, with pSTS and adjacent social-integration hubs as a priori targets.

## Conclusion

In conclusion, through preregistered stacked encoding model analyses applied to naturalistic movieviewing fMRI data from the Healthy Brain Network, we find little support for pervasive early sensory enhancement accounts of autism at the level of feature-resolved cortical representations. Instead, autism-related differences are concentrated in higher-order integration and social-perceptual regions, most notably the pSTS, where high-level visual representations are reduced and feature weighting shifts toward lower-level visual information in a manner that

tracks social symptom severity. At the same time, audio–visual modality preference and unimodal dominance appear largely conserved, arguing against large global shifts in sensory dominance during passive naturalistic viewing. Developmental effects were prominent, lateralized, and modality-congruent, underscoring the need to interpret group differences against ongoing maturation and motion-related sampling constraints in pediatric datasets. Together, these results support mechanistic interpretations related to altered visual feature weighting within social/multisensory networks and demonstrate that naturalistic stimuli coupled with encoding-models can provide a scalable, theory-relevant framework for characterizing how neurodevelopmental conditions may reshape sensory representations in the brain.

## Methods

### Existing Data

This study is based on two pre-registrations (accessible at <http://osf.io/h92gr> and <http://osf.io/47kj6>; summarized in [Table 1](#)) relying on data from releases 1-10 from the openly accessible Healthy Brain Network (HBN) dataset from the Child Mind Institute (Alexander et al., 2017). The HBN dataset offers a large cross-sectional population that is highly phenotypically characterized. The combination of a young population, naturalistic imaging acquisitions, and scales like the social responsiveness score make it an ideal dataset for this analysis.

### Explanation of Existing Data

As additional data is released, a growing body of prior research has analyzed the HBN dataset (Kember et al., 2024; Mihailov et al., 2020; Tong et al., 2024; Cohen et al., 2022; Camacho et al., 2023; Di et al., 2023). To our knowledge, at the time of this writing, there is no existing research using naturalistic fMRI to predict autism diagnostic status and SRS score from this dataset. Some metadata relevant to the research plan including summary statistics, phenotypic data and quality control metrics from the fMRI data were accessed and analyzed prior to the pre-registration of this work. As part of a pilot study to select features and set thresholds and methodological parameters for the pre-registrations, functional data from 54 participants who are neither in the autism nor control groups were accessed. These participants were withheld from all subsequent analyses.

## Behavioral Measures

### Autism diagnosis

Autism diagnosis relied on clinically synthesized consensus DSM-5 diagnoses. These diagnoses were generated by the HBN clinical team based on interviews including the Schedule for Affective Disorders and Schizophrenia—Children’s version (KSADS) and all materials and questionnaires collected during study participation in addition to behavioral observations, family history, previous diagnoses, and history of therapeutic intervention. A limited subset of participants with autism-related traits were administered the Autism Diagnostic Observation Schedule (ADOS). All phenotypic data were accessed in accordance with a data usage agreement from the Child Mind Institute.

### Social Responsiveness Scale (SRS-2)

The SRS is a quantitative measure of autistic traits that is correlated with other commonly used scales including the ADOS and ADI-R (Bölte et al., 2008; Constantino et al., 2003; Constantino, 2007).

### Sensory subset score

While we aim to link neural signatures to targeted behaviorally measured sensory differences in autism, assessments that specifically target such differences, like the Short Sensory Profile (SSP), were not included in the HBN study. To overcome this limitation, a sensory subset score (SSS) was generated from 3 items on the SRS-2 and one item from the Social Communication Questionnaire (SCQ) (Rutter et al., 2003), both of which all participants took part in. While the SRS-2 is a well-

validated measure, the SSS is not psychometrically validated, but may be more specific to sensory-related differences. SSS is composed of the sum of items 20, 42, and 58 of the SRS-2 and item 14 of the SCQ relating to unusual sensory interests, sensory oversensitivities, and concentrating on parts instead of the whole. Specifically, these are: “*I have sensory differences that others find unusual (for example, smelling or looking at things in a different way).*”, “*I am overly sensitive to certain sounds, textures, or smells.*”, “*I concentrate too much on parts of things rather than seeing the whole picture.*”, and “*Has he/she ever seemed to be unusually interested in the sight, feel, sound, taste, or smell of things or people?*”. SSS is a sum of these scores where 0=No/Not True, 1=Sometimes True, 2=Often True, and 3=Almost Always True/Yes with the minimum SSS score being 0 and the max being 12.

## Group-level Confounds

All primary analyses will be conducted at the individual level before demographic confounds are accounted for. Demographics including age, sex, and socioeconomic status along with site-level variability will be included as confounds in all group-level statistical analyses. Family socioeconomic status will be reported as the total score from the Barratt Simplified Measure of Social Status (BSMSS).

## fMRI data

The analysis focuses on two naturalistic movie-viewing fMRI acquisitions (*Despicable Me*, 10 min; *The Present*, 4 min). All fMRI data (TR=0.8s) were acquired from 3 scanner sites in the Greater New York area.

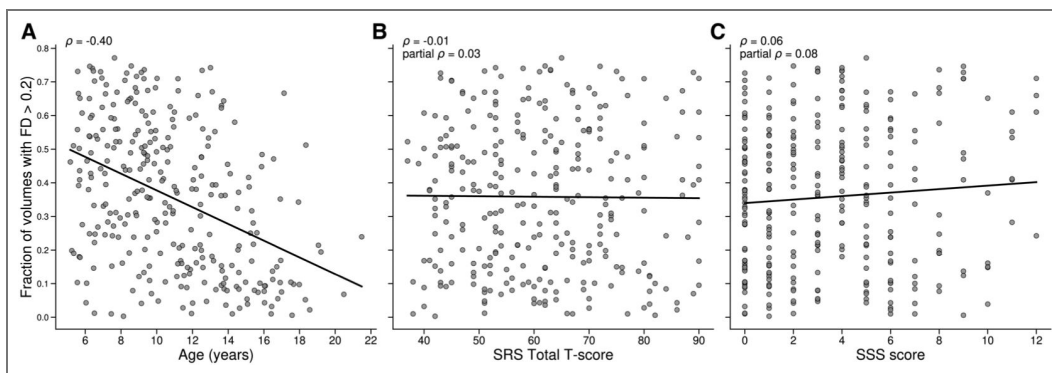
## Quality Control Criteria

Motion-based exclusion criteria in neuroimaging studies of autistic children is an important consideration (Elyounssi et al., 2025 [↗](#)) and quality control criteria have been shown to alter the distribution of included subjects. Specifically, subjects that are younger and have higher ADOS and SRS scores, along with more inattentive, hyperactive/impulsive, and motor control features are more likely to be excluded due to motion (Nebel et al., 2022 [↗](#)). To account for this issue, all results are reported across a sweep of 3 framewise displacement (FD) based functional motion thresholds. Specifically, functional data is excluded if the number of time points in a run with a mean FD > 0.2mm is greater than 40%, 60%, or 80%. By starting with a more stringent threshold and sweeping this across more lenient parameters, we aim to strike a balance between the best quality and more fully characterizing as much of the dataset as possible. In this sample specifically, head motion (fraction of volumes with FD > 0.2 mm) decreased with age (Spearman  $\rho = -0.40$ ), but showed weak linear associations with SRS Total T and SSS ( $\rho = -0.01$  and 0.06, respectively), which remained small after controlling for age (partial  $\rho = 0.03$  and 0.08). As neither SRS nor SSS were correlated with fractional FD, this pattern suggests that age is a primary driver of motion in this cohort and not symptom severity (see Figure 9 [↗](#)).

Quality control is not limited to functional data and is also applied to T1 data. T1 quality is assessed in a semi-automated procedure via a Bayesian DNN brain parcellation method which provides an uncertainty metric highly correlated with T1 structural image quality (McClure et al., 2019 [↗](#)). Images with mean uncertainty  $\geq 0.4$  are flagged for manual inspection of their structural images and subjects with excessive artifacts and irregularities are excluded from all analyses. In addition to these semi-automated approaches, a manual QC evaluation is performed by examining the reconstructed surfaces both alone and overlaid on subject T1 images across all subjects. This entire process is performed “blind”, i.e. all visual inspections are performed without knowledge of which cohort a subject belongs to or any other demographic information.

## Diagnostic Inclusion Criteria

A subset of the total subjects (n=4767 before quality control and before excluding pilot subjects) are included in analyses based on their diagnostic status as determined by the clinician consensus field from the phenotypic and demographic data. Subjects in the autism group are selected if they



**Figure 9. Head motion vs. age and behavioral measures.**

Scatterplots show the fraction of volumes with framewise displacement > 0.2 mm in relation to (A) age, (B) SRS Total T-score, and (C) Sensory Subset Score (SSS). Solid black lines denote linear fits. Reported coefficients are Spearman correlations ( $\rho$ ); “partial  $\rho$ ” controls for age. (A)  $\rho = -0.40$ ; (B)  $\rho = -0.01$ , partial  $\rho = 0.03$ ; (C)  $\rho = 0.06$ , partial  $\rho = 0.08$ . Motion decreases with age and shows only weak associations with SRS/SSS.

have received the DSM-5 label of ‘Autism Spectrum Disorder’ (n=659 before quality control) and non-autistic participants, referred to here as nonASD, are selected if they have the label ‘No Diagnosis Given’ (n=373 before quality control). Individuals with the labels ‘No Diagnosis Given: Incomplete Eval’ and ‘No Diagnosis Given: No Reason Given’ are excluded from all analyses. Any individuals with co-occurring conditions including ADHD are not excluded from the autism group, despite the significant overlap (n=176 autism only, n=483 both autism and ADHD) in order to fully sample the phenotypic space of autism present in the dataset. Additional secondary analyses are performed to compare the autism only group to controls.

## Demographics

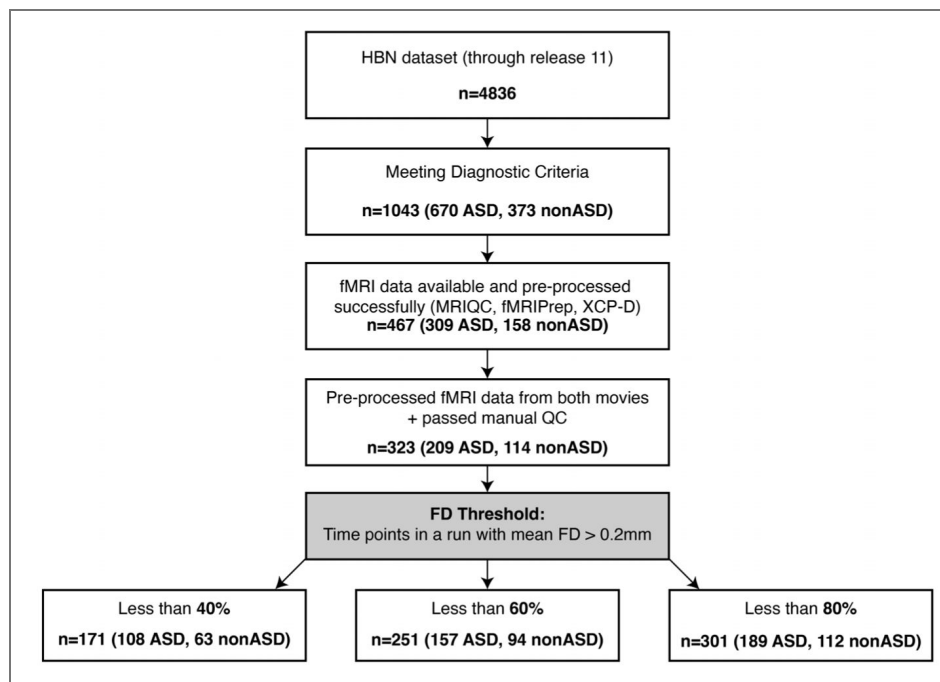
After quality control and filtering (see Figure 10 for a flow chart) at the most stringent FD threshold, 63 non-autistic participants (mean ± SEM age = 10.96 ± 0.46 years; WISC-FSIQ = 110.8 ± 2.05; Barratt Total = 54.4 ± 1.39; SRS Total T = 49.3 ± 0.97; SSS = 1.25 ± 0.21) and 108 autistic participants (mean ± SEM age = 12.76 ± 0.37 years; WISC-FSIQ = 101.4 ± 2.08; Barratt Total = 49.3 ± 1.38; SRS Total T = 68.1 ± 1.10; SSS = 4.71 ± 0.28) were retained. Sex distribution differed significantly between groups ( $\chi^2(1) = 14.23, p = 1.6 \times 10^{-4}$ ), with a higher proportion of males in the autism cohort. Normality (Shapiro-Wilk or D’Agostino’s  $K^2$ ) and homogeneity of variance (Levene’s test) were assessed for each continuous measure within each diagnostic group. Where both normality and equal variance held, Student’s *t*-tests were applied; where variances differed, Welch’s *t*; and for non-normal distributions, two-sided Mann-Whitney *U*-tests were used. Sex differences were evaluated via  $\chi^2$ , with Fisher’s exact test substituted if any cell count fell below five. To control the family-wise error rate over six univariate comparisons, *p*-values were adjusted via the Benjamini-Hochberg false discovery rate ( $\alpha = 0.05$ ). All demographic measures differed significantly between non-autistic and autistic groups after FDR correction, but a multivariate logistic regression (diagnosis ~ Age + FSIQ + Barratt + SRS + SSS + Sex) confirmed that, when entered simultaneously, only SRS ( $z = -3.84, p = 1.2 \times 10^{-4}$ ) and sex ( $z = 2.70, p = 6.8 \times 10^{-3}$ ) remained significant independent predictors of diagnostic status.

Group	Statistic	Age	WISC FSIQ	Barratt Total	SRS Total T	SSS	Sex
Autism	Mean	12.76	101.43	49.30	68.14	4.71	
	SEM	0.37	2.08	1.38	1.10	0.28	
	Min	5.25	51.00	9.00	44.00	0.00	
	Max	21.48	147.00	66.00	90.00	12.00	
	N	108	84	107	103	103	90 M, 18 F
Control	Mean	10.96	110.76	54.38	49.32	1.25	
	SEM	0.46	2.05	1.39	0.97	0.21	
	Min	5.35	81.00	15.50	38.00	0.00	
	Max	20.49	145.00	66.00	73.00	8.00	
	N	63	51	63	62	61	35 M, 28 F
<b>p-value</b>		<b>0.004</b>	<b>0.004</b>	<b>0.023</b>	<b>8.51e-19</b>	<b>7.55e-14</b>	<b>3.23e-04</b>

**Table 5. Descriptive statistics of demographic and clinical variables by diagnostic group.** Continuous measures (Age, WISC-FSIQ, Barratt Total, SRS Total T, SSS) are reported along with their mean, SEM, minimum-maximum range in parentheses. “N” indicates the number of non-missing observations. Sex shows counts of males and females. P-values are FDR-corrected (Benjamini-Hochberg,  $\alpha=0.05$ ).

## fMRI Preprocessing

fMRI data from all included subjects was preprocessed using fMRIPrep version 23.2.0 (Esteban et al., 2019). This version of fMRIPrep includes MSMSulc cortical surface registration, allowing for improved surface alignment compared to previous versions of fMRIPrep. CIFTI2 outputs (combined cortical surface and subcortical volume) were obtained in fsLR 32k space. All subsequent analyses were performed using CIFTI images to better respect the geometry of brain structures of interest and prevent issues such as smoothing across cortical folds. The first 10



**Figure 10.** A flow chart illustrating how participants were selected for the final sample including the number that were excluded at each stage.

timepoints from each run were discarded. Further preprocessing was performed using XCP-D version 0.10.1 (Mehta et al., 2024 [↗](#)). Specific steps included: 4mm FWHM spatial smoothing, despiking, 0.1 Hz low-pass filtering, confound regression of: six motion estimates and their derivatives, five white matter, and five CSF ACompCor components, denoising and detrending.

## Parcellation Scheme

The Glasser Multi-Modal Parcellation (MMP) (Glasser et al., 2016 [↗](#)) was used to select regions of interest for all analyses. Regions were selected from auditory regions and across the dorsal, ventral, and lateral visual streams.

## Feature Selection

Movie features include automatically extracted and hand-labeled low- and high-level audio and visual classes. These features were selected based on their performance on an independent set of 54 pilot subjects that were excluded from all subsequent analyses. Low-level auditory features were extracted from the audio track of the movie files and include perceptual loudness extracted via pyloudnorm (Steinmetz, 2023 [↗](#)) and the first 5 principal components of the cochleagram extracted via pycochleagram (McDermott, 2018 [↗](#)). AudioSet (Gemmeke et al., 2017 [↗](#)) labels comprise the high-level auditory features and were extracted using YAMNet. Instead of a 521-dimensional vector of all AudioSet classes, classes were combined across the highest level of the AudioSet ontology yielding an 8-dimensional feature consisting of: 'Human Sounds, Speech', 'Human Sounds, Non-Speech', 'Animal Sounds', 'Music', 'Natural Sounds', 'Sounds of Things', 'Source Ambiguous Sounds', and 'Channel, Environment, Background'.

Low-level visual features include mean perceptual brightness, extracted from the RGB values from each movie frame and mean motion, calculated as the mean of 2,139 motion energy filters from pymoten (Huth et al., 2012 [↗](#); Naselaris et al., 2011 [↗](#); Nunez-Elizalde et al., 2022 [↗](#)). High-level visual features here include hand-coded labels for the presence of faces and bodies obtained from the EmoCodes dataset (Camacho et al., 2022 [↗](#)). While many more features could be extracted and included in the models, here we limit features to this number as encoding models from the pilot study with greater numbers of features resulted in overfitting.

## Encoding Models

The hypotheses in this study are tested via encoding models, specifically looking at model performance, unique variance explained, and model weights. Individual subject grayordinate-wise stacked regression (Lin et al., 2024 [↗](#)) is used, training a separate encoding model for each grayordinate to predict withheld brain activation for each subject. Three different stacked encoding models are employed here: a low- and high-level audio model with only audio features, a low- and high-level visual model with only visual features, and an audio and visual model with all audio and visual features (the audiovisual stacked model is depicted in Figure 12 [↗](#)). For each model and subject, the significance of each grayordinate is tested via a null model by repeatedly temporally permuting the order of observations and retraining and testing the models. 1,000 iterations are performed for each subject. FDR correction is applied to account for multiple comparisons for modeling and testing a separate model for each grayordinate across the brain. Model weights and performance are interpreted by pooling data across significant grayordinates within parcels of interest. The encoding models use 5-fold cross-validation, training on 80% of the movie data and testing on the remaining 20% across successive folds. To account for the hemodynamic response, all features are convolved with a canonical double gamma HRF function. Features are z-scored to reduce the effect of scaling on model weights. Accuracy is quantified for each region as the mean coefficient of determination ( $R^2$ ) and mean noise ceiling normalized  $R^2$ . Spearman-Brown corrected split-half noise ceilings are calculated to estimate the upper bound of possible performance for each brain parcel. Unique low-level model  $R^2$  is calculated by subtracting the high-level model  $R^2$  from the stacked encoding model  $R^2$ . For this calculation, when high-level  $R^2$  is less than zero, it is clipped to zero, resulting in low-level  $R^2 = \text{stacked } R^2$ . When high-level  $R^2$  exceeds stacked  $R^2$  it is set to stacked  $R^2$ , resulting in low-level  $R^2 = 0$ . Model weight difference is

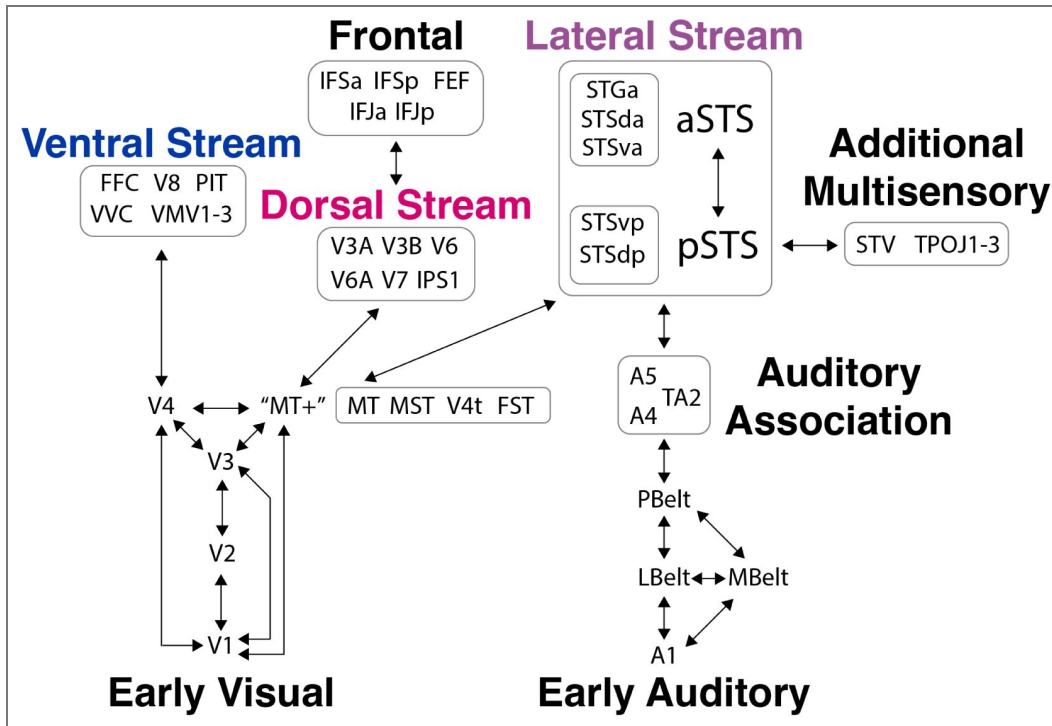
calculated by subtracting the high-level stacked encoding model weight  $\alpha_H$  from the low-level weight  $\alpha_L$  (for the audiovisual model, the visual weight  $\alpha_V$  is subtracted from the audio weight  $\alpha_A$ ) resulting in a value [-1,1] where 1 indicates a maximum high-level preference, -1 indicates a maximum low-level preference and 0 indicates no preference between features.

## Statistical Analysis

Linear mixed-effect models are used for significance testing for all hypotheses. While testing the relationship between model performance ( $R^2$ ) and diagnostic group membership (control vs autism), age, sex, socioeconomic status (SES), and site-level variability are accounted for as covariates. IQ is not included as a covariate in the model due to too many participants with missing data, but a multivariate logistic regression including all covariates and WISC full-scale IQ (the IQ assessment with the largest coverage across the included participants) confirmed that only SRS and sex remained significant independent predictors of diagnostic status.  $R^2$ , the coefficient of determination, here represents model fit and serves as the dependent variable. Fixed effects include group membership (binary: 0 or 1), age (continuous, in years), sex (binary: 0 = Female, 1 = Male), and SES (continuous). Site (categorical, across 3 sites) is included as a random intercept to account for unmeasured heterogeneity across sites, allowing the model to adjust for clustering effects. A generalized linear mixed-effects model (GLMM) is fit using a logit link function. The model is implemented in Python using the *pymr4* library (Jolly, 2018 [DOI](#)). Fixed-effect estimates are tested for statistical significance, and the variance of the random effect is assessed to quantify site-level heterogeneity. This same procedure is applied to test the other dependent variables in the low- vs. high-level and audio vs. visual hypotheses: unique explained variance and model weight difference. To test for a relationship with autism severity and sensory symptoms, separate mixed-effects models are employed for each of these metrics, but with the behavioral measures SRS and SSS each included as independent variables of interest in place of diagnostic group membership.

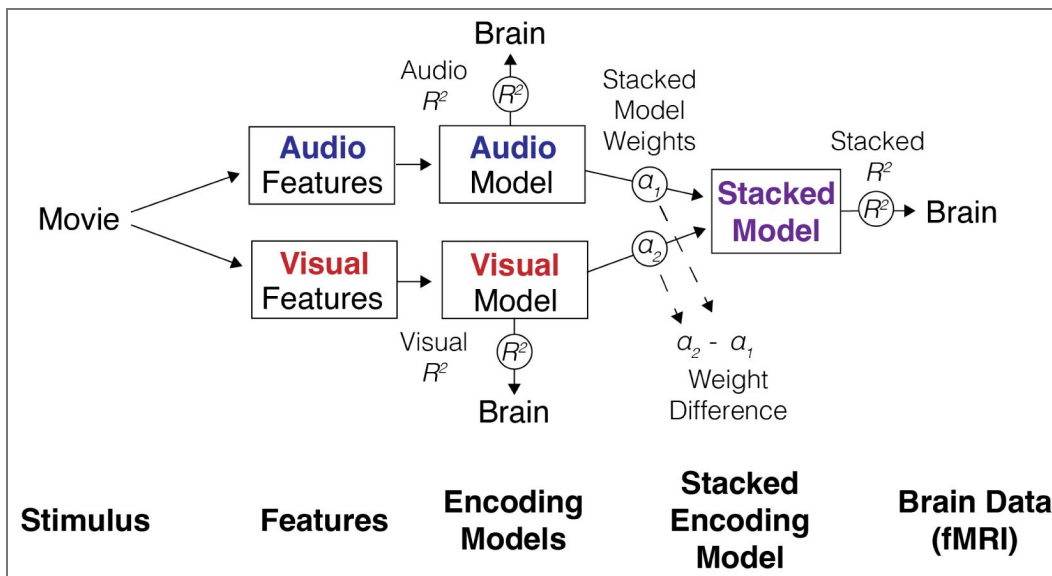
**Figure 11. Brain regions of interest with labels selected from the MMP span the visual and auditory systems and the dorsal, ventral, and lateral streams.**

The schematic shows all perceptual regions of interest grouped by their classification from the MMP and with perceptual streams labeled and some simplified connections illustrated. Note that the actual connectivity between these regions is known to be far more complex, for vision, see Felleman and Van Essen (1991) [link](#).

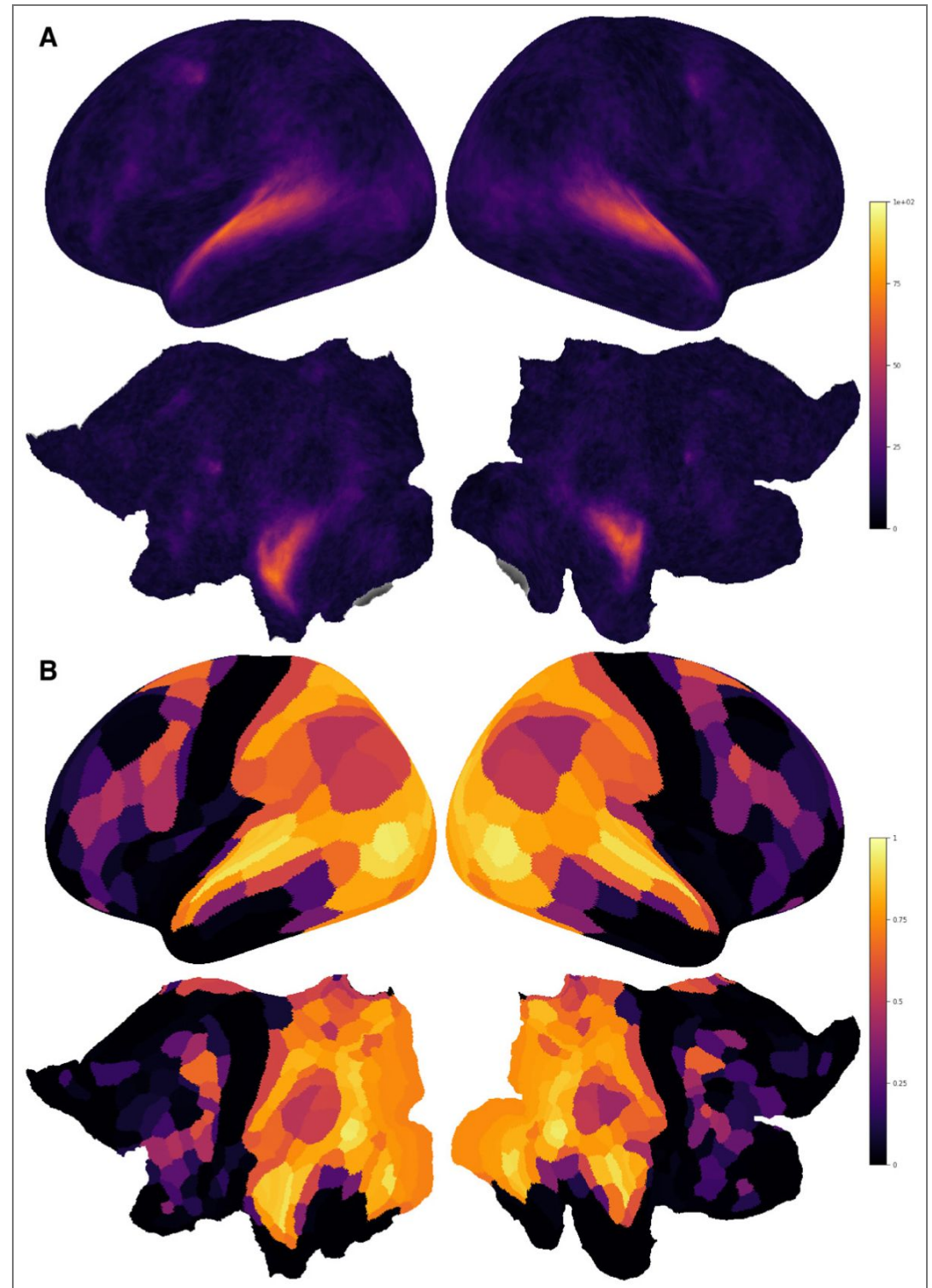


**Figure 12. Schematic diagram illustrating the stacked encoding model approach used to relate movie stimulus features to brain activation measured via fMRI in the case of the audio and visual stacked model.**

From left to right: audio and visual features are extracted from a naturalistic movie stimulus. Each feature set independently informs its own ridge regression encoding model, which predicts brain activation ( $R^2$ ) on held-out data. The audio and visual ridge regression models are subsequently combined into a stacked regression model, assigning each model a weight ( $\alpha_1, \alpha_2$ ). The performance of the stacked model is measured by predicting overall brain activation ( $R^2$ ). The difference between stacked model weights ( $\alpha_2 - \alpha_1$ ) quantifies the relative perceptual preference toward audio or visual features.



## Figure supplements



**Figure 2—figure supplement 1. Overall model performance across the cortex.** (A) Whole-brain plot of the percentage of subjects with a significant grayordinate at each region. The significance of each grayordinate was tested via a null model by repeatedly temporally permuting the order of observations and retraining and testing the models over 1,000 permutations. (B) The Spearman-Brown corrected split-half noise ceilings for each MMP parcel.

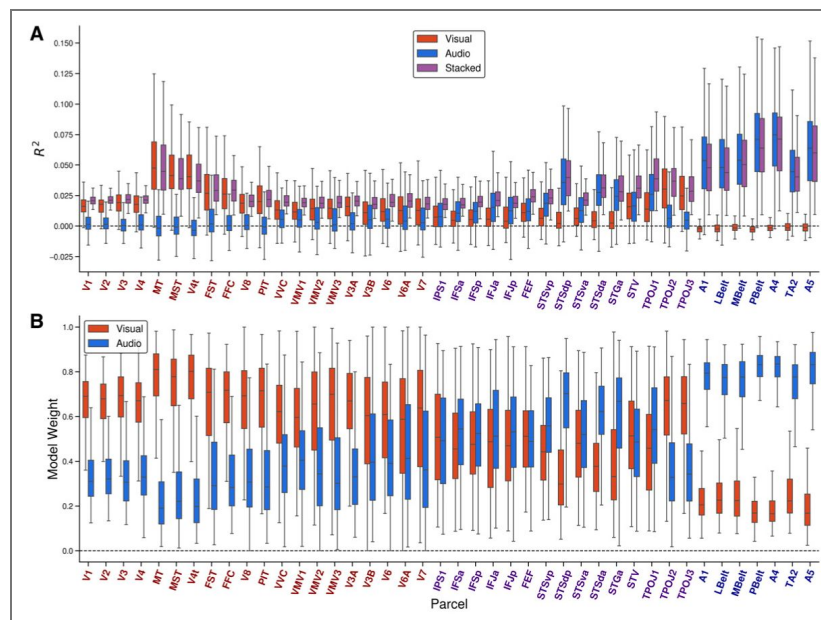


Figure 2—figure supplement 2.

(A) Model performance ( $R^2$ ) for the visual (red), audio (blue), and stacked (purple) encoding models across perceptual regions of interest for all included participants (both ASD and nonASD). (B) Visual (red) and audio (blue) model weights from the stacked encoding model.

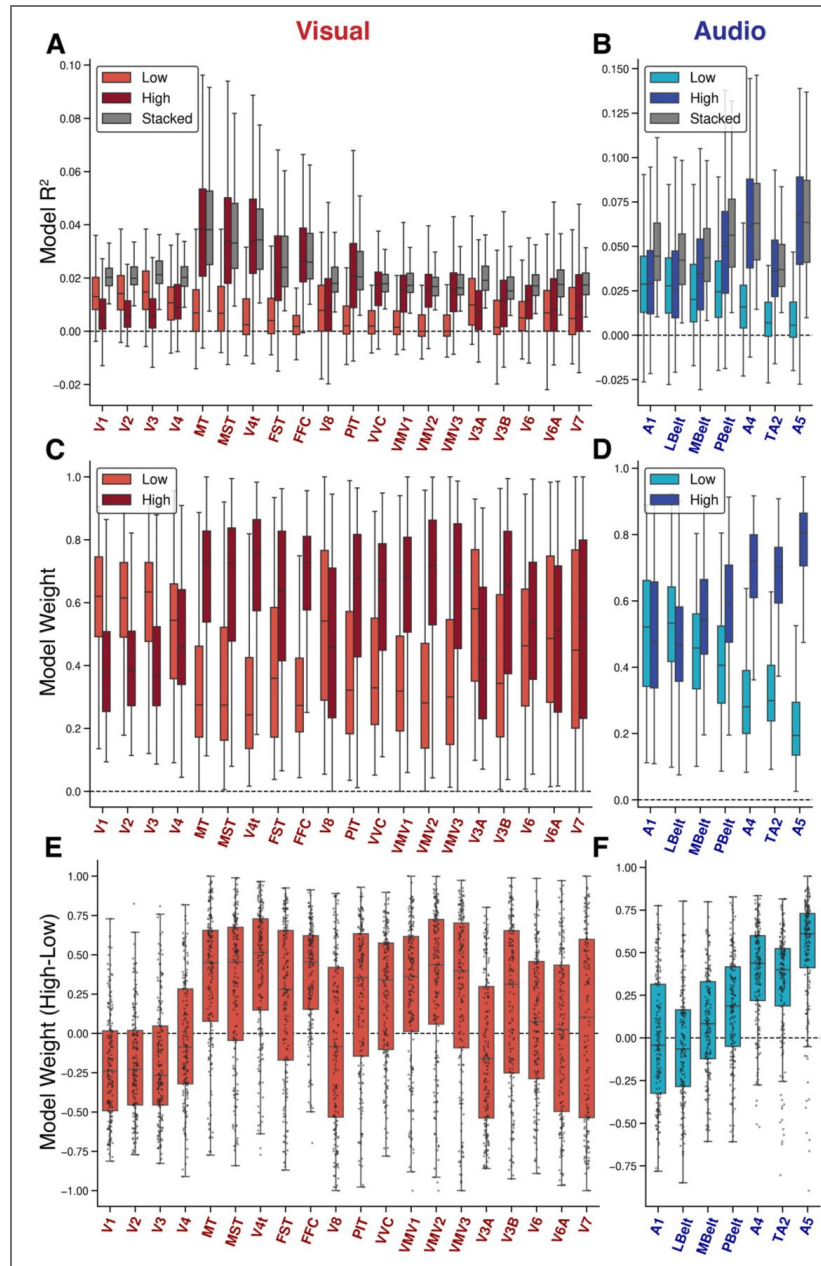
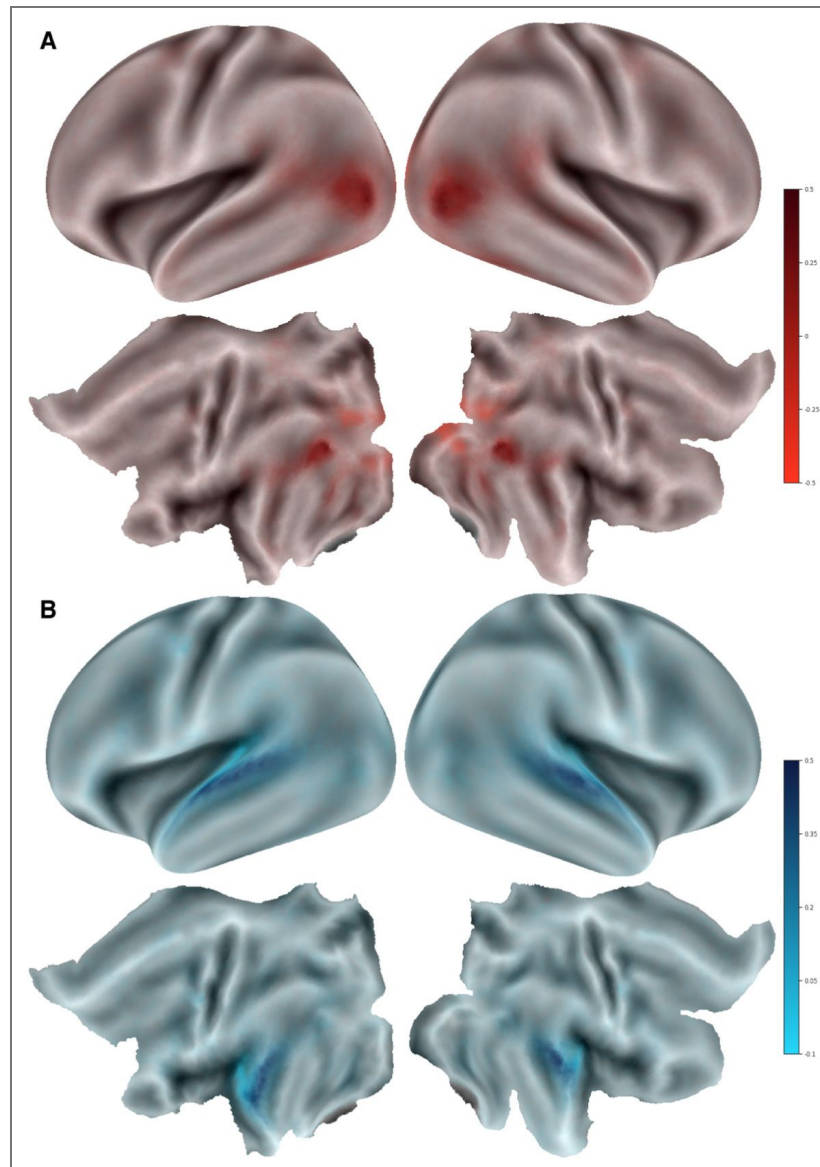


Figure 2—figure supplement 3.

(A) Model performance ( $R^2$ ) for the visual low- (light red) and high-level (dark red) and stacked (grey) encoding models across all nonASD and ASD participants. (A) Corresponding  $R^2$  for audio models. (C) Model weights from the stacked visual encoding model. (D) Model weights from the stacked audio encoding model. (E), (F) High- vs. low-level visual and audio perceptual preferences ( $W_H - W_L$ ) are calculated by taking the difference of high- and low-level weights (shown in B and C). Perceptual preference is in the range of -1 to 1, as the stacked encoding model weights range from 0 to 1. Positive values here indicate a high-level preference, negative values indicate a low-level preference, and values around zero indicate no preference.



**Figure 2—figure supplement 4.**

(A) Whole-brain grayordinate-wise plot of mean high- vs. low-level perceptual preference across all participants (ASD and nonASD). (B) The same corresponding plot but from the audio stacked encoding model.

Figure 3—figure supplement 1.

(A) Box plot of low-level audio encoding model explained variance ( $R^2$ ) for ASD (white) and nonASD (gray) groups across all perceptual ROIs. (B) Box plot of the unique explained variance ( $R_u^2$ ). The results correspond to the 40% FDR threshold. Boxes annotated with circles indicate an initially significant difference between groups that did not survive FDR correction. Boxes show the quartiles of the dataset, and whiskers show the distribution, with the exception of outliers. Each dot is the mean  $R^2$  from a single subject from all statistically significant grayordinates within each ROI.

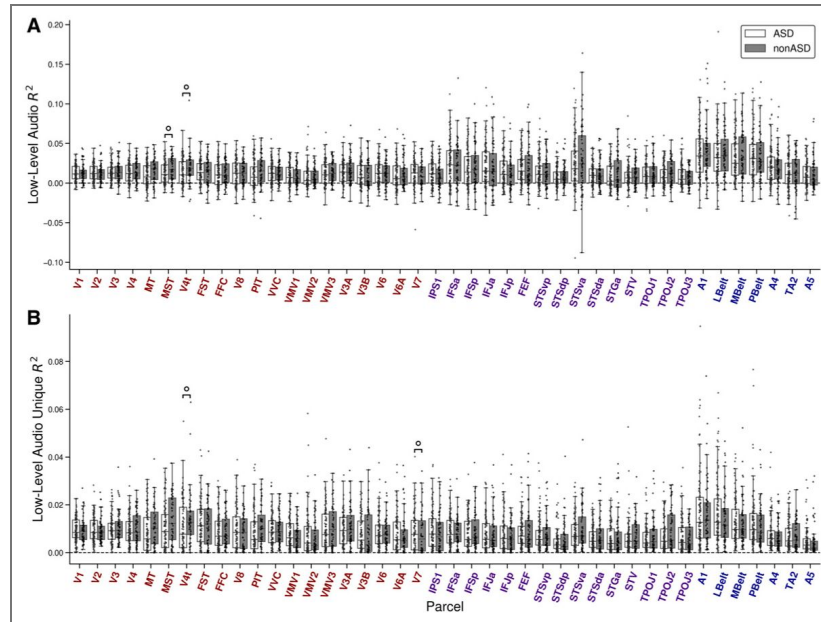


Figure 3—figure supplement 2.

(A) Box plot of low-level visual encoding model  $R^2$  for ASD (white) and nonASD (gray) groups across all perceptual ROIs. (B) Corresponding box plot of  $R_u^2$ . Results correspond to the 40% FDR threshold. Boxes annotated with an asterisk indicate a significant group difference (FDR  $p < 0.05$ ), while a circle indicates an initially significant difference between groups that did not survive FDR correction. Boxes show the quartiles of the dataset and whiskers show the distribution with the exception of outliers. Each dot is the mean  $R^2$  from a single subject from all statistically significant grayordinates within each ROI.

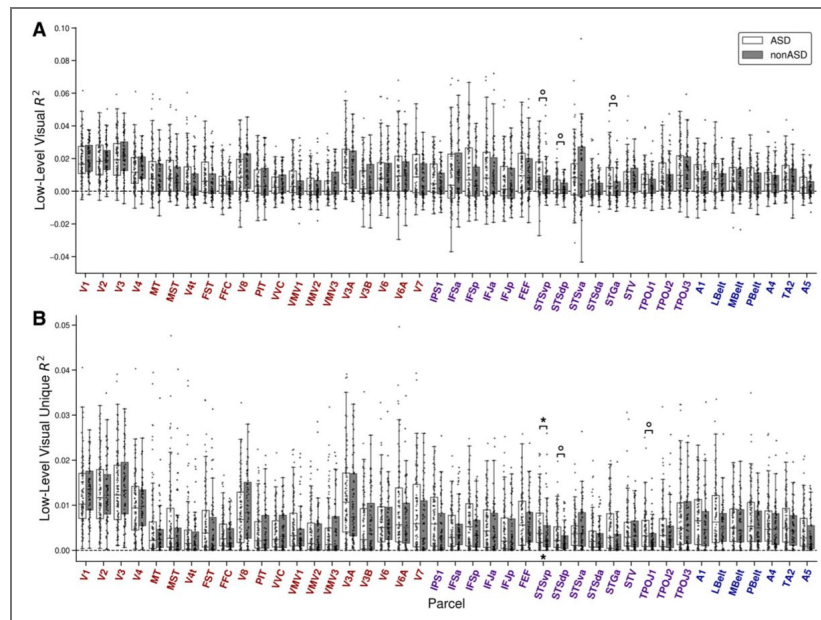


Figure 3—figure supplement 3.

(A) Box plot of high-level audio encoding model  $R^2$  for ASD (white) and nonASD (gray) groups across all perceptual ROIs. (B) Corresponding box plot of  $R_U^2$ . Results correspond to the 40% FD threshold. Boxes annotated with an asterisk indicate a significant group difference (FDR  $p < 0.05$ ), while a circle indicates an initially significant difference between groups that did not survive FDR correction. Boxes show the quartiles of the dataset and whiskers show the distribution with the exception of outliers. Each dot is the mean  $R^2$  from a single subject from all statistically significant grayordinates within each ROI.

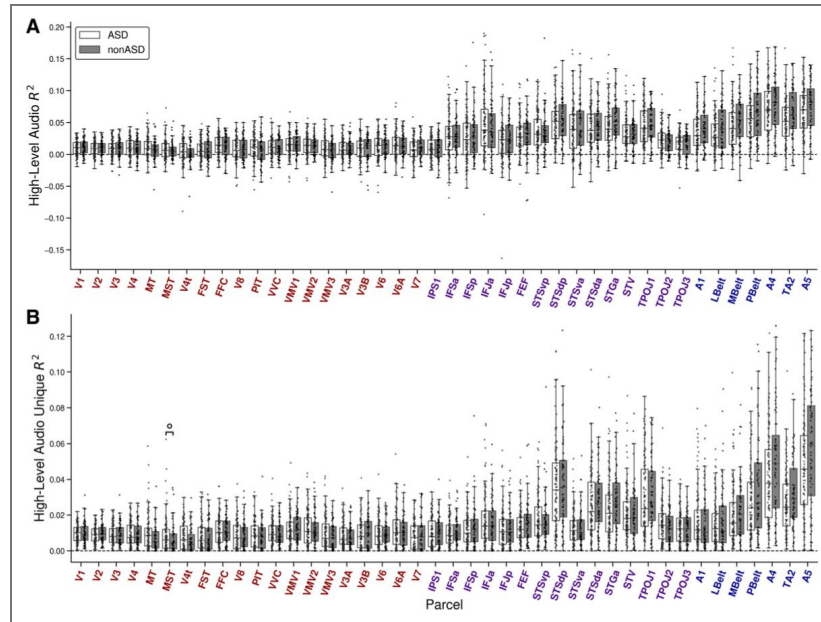
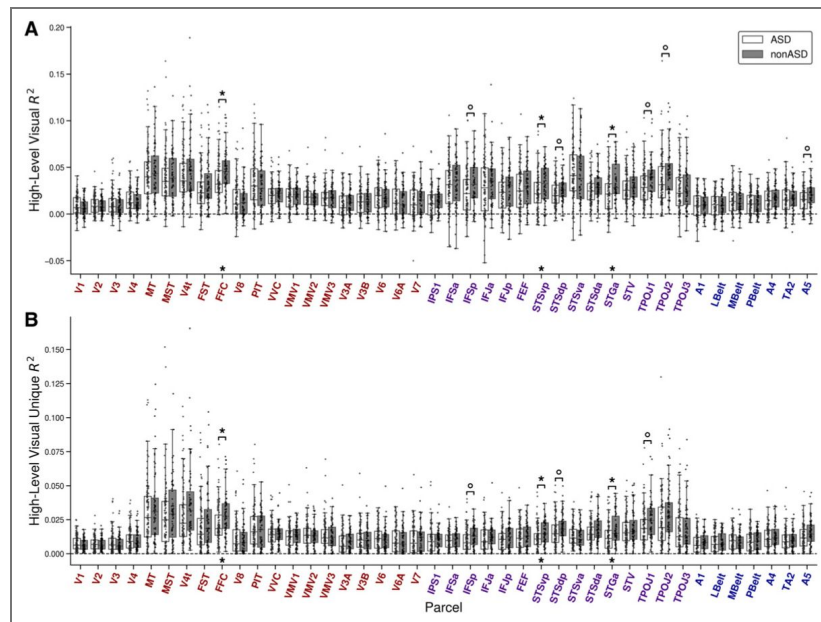


Figure 3—figure supplement 4.

(A) Box plot of high-level visual encoding model  $R^2$  for ASD (white) and nonASD (gray) groups across all perceptual ROIs. (B) Corresponding box plot of  $R_U^2$ . Results correspond to the 40% FD threshold. Boxes annotated with an asterisk indicate a significant group difference (FDR  $p < 0.05$ ), while a circle indicates an initially significant difference between groups that did not survive FDR correction. Boxes show the quartiles of the dataset and whiskers show the distribution with the exception of outliers. Each dot is the mean  $R^2$  from a single subject from all statistically significant grayordinates within each ROI.



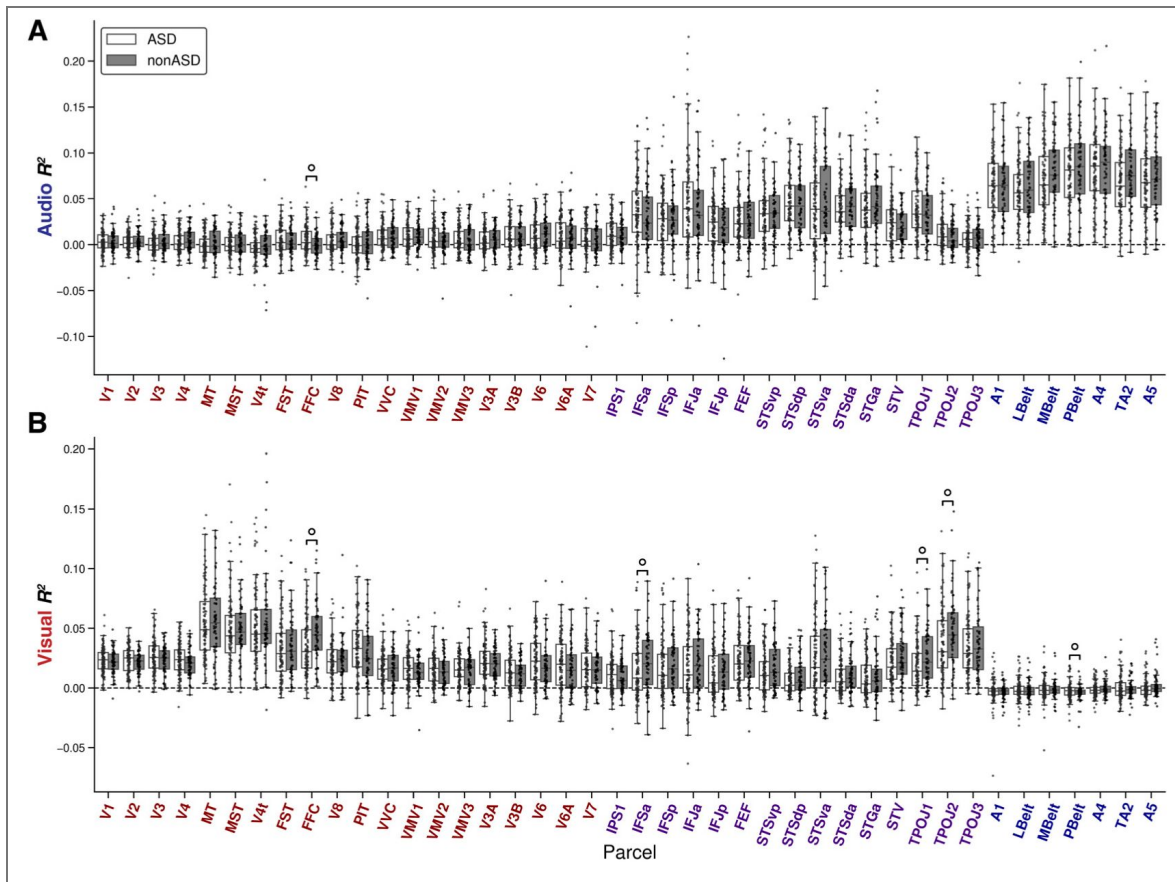


Figure 7—figure supplement 1.

(A) Box plot of audio encoding model  $R_U^2$  for ASD (white) and nonASD (gray) groups across all perceptual ROIs. (B) The same but for visual  $R_U^2$ . Results correspond with the 40% FD threshold. Boxes annotated with circles indicate a difference between groups that did not survive FDR correction (uncorrected  $p < 0.05$ ,  $FDR q > 0.05$ ). Boxes show the quartiles of the dataset and whiskers show the distribution with the exception of outliers. Each dot is the mean value from a single subject from all statistically significant grayordinates within each ROI.

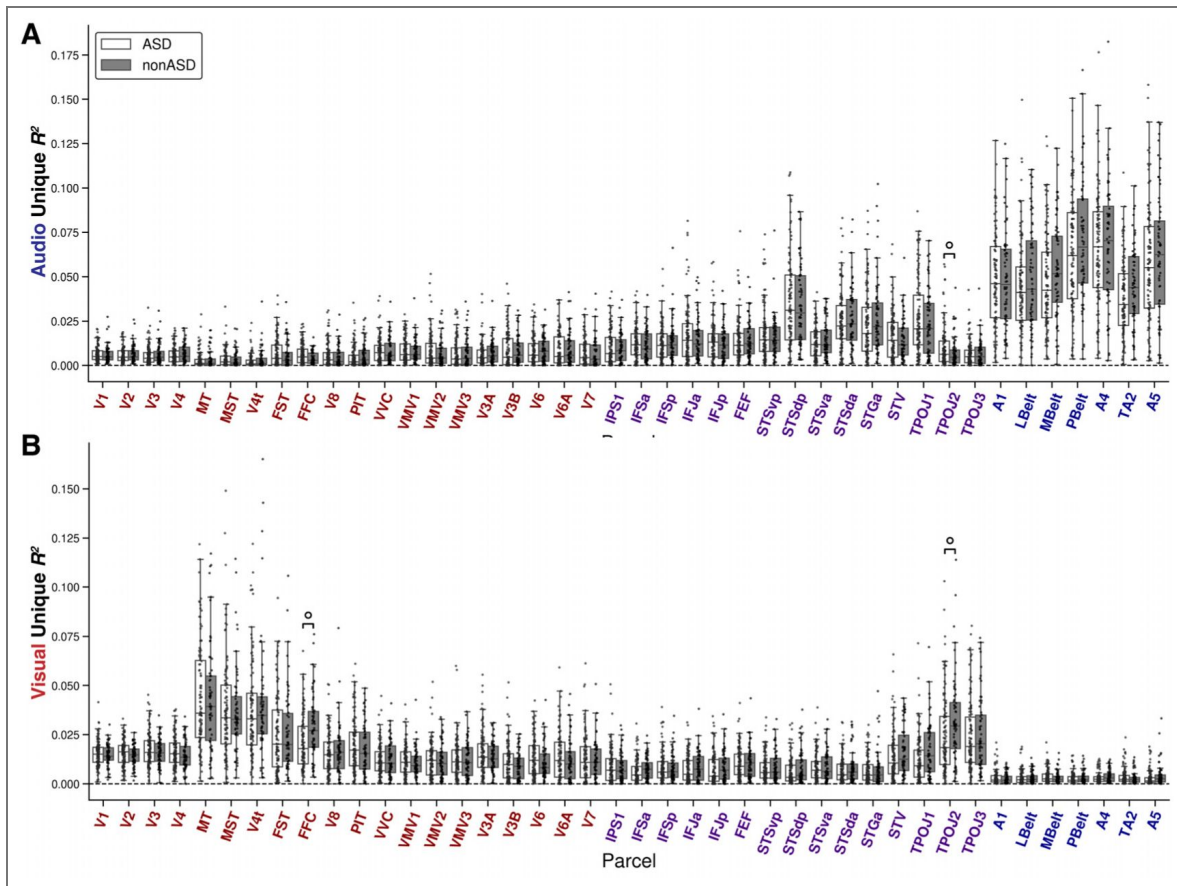
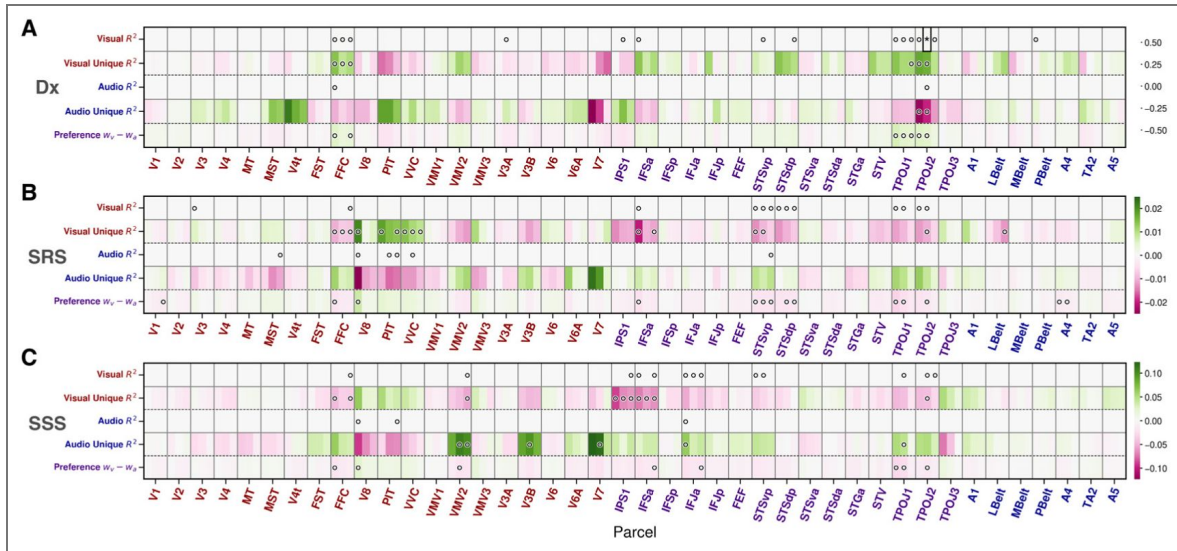


Figure 7—figure supplement 2.

(A) Box plot of audio encoding model  $R^2$  for ASD (white) and nonASD (gray) groups across all perceptual ROIs. (B) The same but for visual  $R^2$ . These results correspond with the 40% FD threshold. Boxes annotated with circles indicate a difference between groups that did not survive FDR correction (uncorrected  $p < 0.05$ , FDR  $q > 0.05$ ). Boxes show the quartiles of the dataset and whiskers show the distribution with the exception of outliers. Each dot is the mean value from a single subject from all statistically significant grayordinates within each ROI.

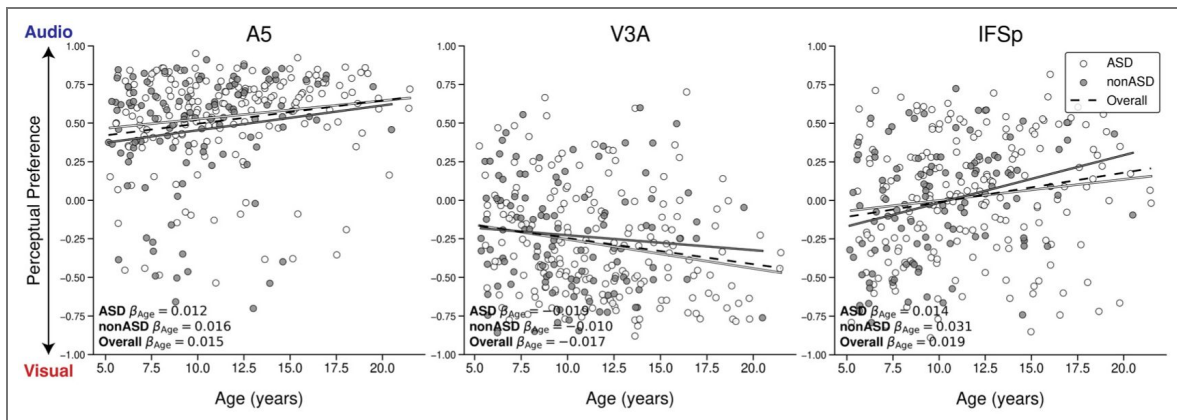
**Figure 7—figure supplement 3.** Heatmaps of fixed-effect coefficients for (A) diagnostic group (Dx: nonASD vs. ASD), (B) SRS Total T-Score and (C) sensory subset score (SSS), across 40%, 60% and 80% FD thresholds (left to right within each cortical parcel column).

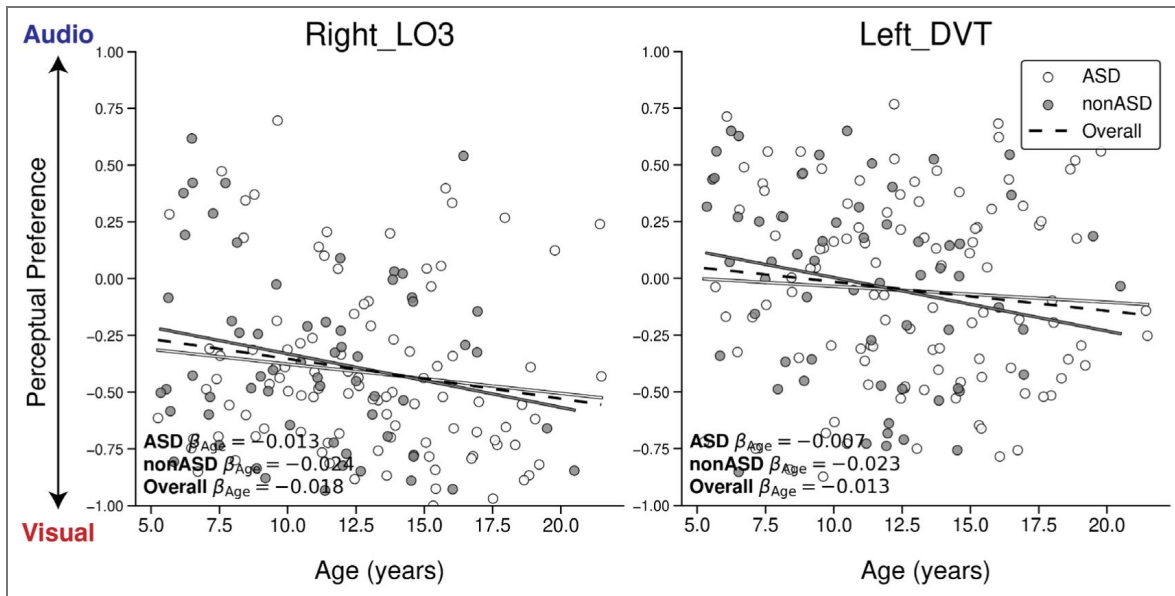
Within each of the three horizontal panels, rows denote encoding model-derived metrics (visual  $R^2$  and  $R_U^2$ , audio  $R^2$  and  $R_U^2$ , and their preference index ( $W_V - W_A$ )) and columns denote Glasser MMP ROIs ordered left to right from early visual areas through association cortices followed by auditory areas. Color indicates the magnitude and sign of the coefficient (pink=negative effect with ASD>nonASD; green=positive effect with ASD<nonASD). Asterisks mark FDR-corrected significance at  $q < 0.05$ ; open circles mark uncorrected  $p < 0.05$ . The visual modality encoding models are labeled in red and the auditory in blue.



**Figure 8—figure supplement 1.** Scatter plots showing the relationship between age and perceptual preference in three example cortical regions (A5, V3A, and IFSp) where significant effects were observed across all participants.

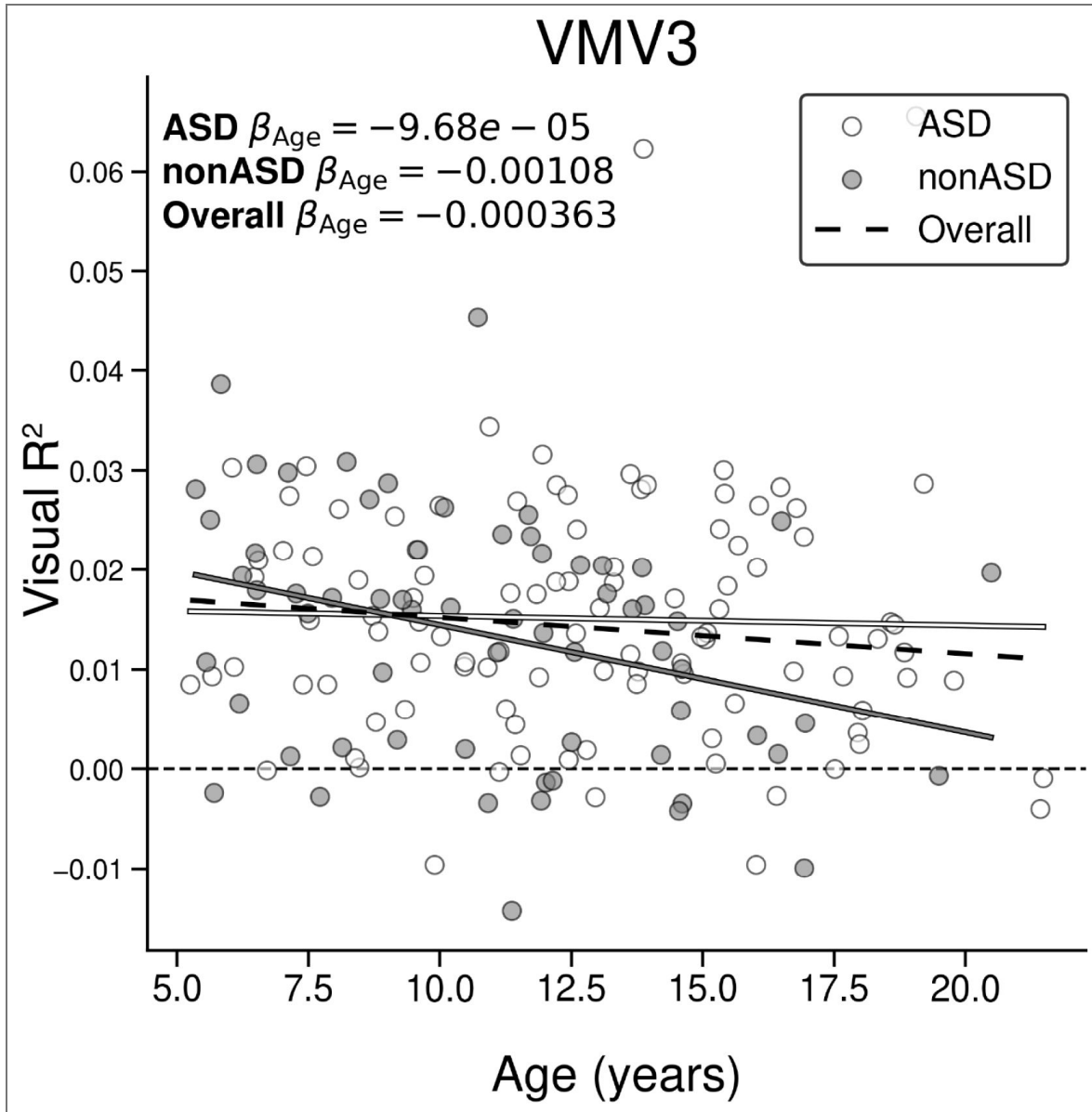
Perceptual preference values above zero indicate an auditory preference, and values below zero indicate a visual preference. Although no significant age-by-diagnosis interactions are displayed here, autistic (ASD; white) and non-autistic (nonASD; gray) groups are colored separately for clarity. Lines show model fits for each group colored correspondingly, with estimated age-related slopes ( $\beta_{Age}$ ) reported for ASD, nonASD, and the overall sample (black dashed line).





**Figure 8—figure supplement 2.** Scatter plots showing the relationship between age and perceptual preference in two example lateralized cortical regions, Right LO3 (a visual region between early visual cortex and MT+) and Left DVT (Dorsal Visual Transitional area; a region located on the posterior bank of the parieto-occipital sulcus), where significant effects were observed across all participants at the whole-brain level.

Perceptual preference values above zero indicate an auditory preference, and values below zero indicate a visual preference. Although no significant age-by-diagnosis interactions are displayed here, autistic (ASD; white) and non-autistic (nonASD; gray) groups are colored separately for clarity. Lines show model fits for each group, with estimated age-related slopes ( $\beta_{Age}$ ) reported for ASD, nonASD, and the overall sample (black dashed line).



**Figure 8—figure supplement 3.** Scatter plot showing the relationship between age and visual  $R^2$  in perceptual region VMV3 where a significant age:diagnosis interaction was observed.

Autistic (ASD; white) and non-autistic (nonASD; gray) diagnostic groups are displayed. Lines show model fits for each group, with estimated age-related slopes ( $\beta_{Age}$ ) reported for ASD, nonASD, and the overall sample (black dashed line).

## Data availability

The current manuscript is a computational study, so no data have been generated for this manuscript.

## Acknowledgements

We would like to thank Cat Camacho along with the members of the Senseable Intelligence Group for their insightful discussions.

## Additional information

### Funding

JM, YC, and SSG were partially supported by NIH P4 EB019936, by gifts from Schmidt Futures and Citadel founder and CEO Ken Griffin, and by the Lann and Chris Woehrle Psychiatric Fund at the McGovern Institute for Brain Research at MIT. JM was supported by the NIH-NIDCD T32 (5T32DC000038-29).

### Funding

Funder	Grant reference number	Author
National Institutes of Health	5T32DC000038-29	Jeff Mentch
National Institutes of Health	NIH P4 EB019936	Yibei Chen Satrajit S Ghosh Jeff Mentch

### Author ORCID iDs

Satrajit S Ghosh:  <https://orcid.org/0000-0002-5312-6729>

## References

- Abrams DA, Lynch CJ, Cheng KM, Phillips J, Supekar K, Ryali S, Uddin LQ, Menon V (2013) Underconnectivity between voice-selective cortex and reward circuitry in children with autism. *Proceedings of the National Academy of Sciences* **110**:12060-12065 <https://doi.org/10.1073/pnas.1302982110> | PubMed
- Alaerts K, Nayar K, Kelly C, Raithe J, Milham MP, Di Martino A (2015) Age-related changes in intrinsic function of the superior temporal sulcus in autism spectrum disorders. *Social cognitive and affective neuroscience* **10**:1413-1423 <https://doi.org/10.1093/scan/nsv029> | PubMed
- Alaerts K, Woolley DG, Steyaert J, Di Martino A, Swinnen SP, Wenderoth N (2014) Underconnectivity of the superior temporal sulcus predicts emotion recognition deficits in autism. *Social cognitive and affective neuroscience* **9**:1589-1600 <https://doi.org/10.1093/scan/nst156> | PubMed
- Alexander LM, Escalera J, Ai L, Andreotti C, Febre K, Mangone A, Vega-Potler N, Langer N, Alexander A, Kovacs M, et al. (2017) An open resource for transdiagnostic research in pediatric mental health and learning disorders. *Sci Data* **4** <https://doi.org/10.1038/sdata.2017.181> | PubMed
- Balasco L, Provenzano G, Bozzi Y (2020) Sensory abnormalities in autism spectrum disorders: a focus on the tactile domain, from genetic mouse models to the clinic. *Frontiers in psychiatry* **10**:1016 <https://doi.org/10.3389/fpsy.2019.01016> | PubMed
- Ben-Sasson A, Cermak SA, Orsmond GI, Tager-Flusberg H, Carter AS, Kadlec MB, Dunn W (2007) Extreme sensory modulation behaviors in toddlers with autism spectrum disorders. *The American Journal of Occupational Therapy* **61**:584-592 <https://doi.org/10.5014/ajot.61.5.584> | PubMed
- Bennetto L, Kushner ES, Hyman SL (2007) Olfaction and taste processing in autism. *Biological psychiatry* **62**:1015-1021 <https://doi.org/10.1016/j.biopsych.2007.04.019> | PubMed

- Bolton TAW**, Freitas LGA, Jochaut D, Giraud AL, Van De Ville D (2020) Neural responses in autism during movie watching: Inter-individual response variability co-varies with symptomatology. *Neuroimage* **216**:116571 <https://doi.org/10.1016/j.neuroimage.2020.116571> | PubMed
- Bolton TAW**, Jochaut D, Giraud AL, Van De Ville D (2018) Brain dynamics in ASD during movie-watching show idiosyncratic functional integration and segregation. *Hum Brain Mapp* **39**:2391-2404 <https://doi.org/10.1002/hbm.24009> | PubMed
- Brock J** (2012) Alternative Bayesian accounts of autistic perception: comment on Pellicano and Burr. *Trends Cogn Sci* **16**:573-4 <https://doi.org/10.1016/j.tics.2012.10.005> | PubMed
- Burr D**, Gori M (2012) Multisensory Integration Develops Late in Humans. In: Murray MM, Wallace MT (Eds). *The Neural Bases of Multisensory Processes Frontiers in Neuroscience* Boca Raton, FL: CRC Press/Taylor & Francis. <https://www.ncbi.nlm.nih.gov/books/NBK92864/> | PubMed
- Byrge L**, Dubois J, Tyszka JM, Adolphs R, Kennedy DP (2015) Idiosyncratic brain activation patterns are associated with poor social comprehension in autism. *J Neurosci* **35**:5837-5850 <https://doi.org/10.1523/jneurosci.5182-14.2015> | PubMed
- Bölte S**, Poustka F, Constantino JN (2008) Assessing autistic traits: cross-cultural validation of the social responsiveness scale (SRS). *Autism Res* **1**:354-363 <https://doi.org/10.1002/aur.49> | PubMed
- Calvert GA**, Campbell R, Brammer MJ (2000) Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. *Curr Biol* **10**:649-657 [https://doi.org/10.1016/s0960-9822\(00\)00513-3](https://doi.org/10.1016/s0960-9822(00)00513-3) | PubMed
- Calvert GA**, Hansen PC, Iversen SD, Brammer MJ (2001) Detection of audio-visual integration sites in humans by application of electrophysiological criteria to the BOLD effect. *Neuroimage* **14**:427-438 <https://doi.org/10.1006/nimg.2001.0812> | PubMed
- Camacho MC**, Nielsen AN, Balsler D, Furtado E, Steinberger DC, Fruchtmann L, Culver JP, Sylvester CM, Barch DM (2023) Large-scale encoding of emotion concepts becomes increasingly similar between individuals from childhood to adolescence. *Nat Neurosci* <https://doi.org/10.1038/s41593-023-01358-9> | PubMed
- Camacho MC**, Williams EM, Balsler D, Kamojjala R, Sekar N, Steinberger D, Yarlagadda S, Perlman SB, Barch DM (2022) EmoCodes: A standardized coding system for Socio-emotional content in complex video stimuli. *Affect Sci* **3**:168-181 <https://doi.org/10.1007/s42761-021-00100-7> | PubMed
- Cary E**, Pacheco D, Kaplan-Kahn E, McKernan E, Matsuba E, Prieve B, Russo N (2024) Brain signatures of early and late neural measures of auditory habituation and discrimination in autism and their relationship to autistic traits and sensory overresponsivity. *Journal of Autism and Developmental Disorders* **54**:1344-1360 <https://doi.org/10.1007/s10803-022-05866-8> | PubMed
- Cascio CJ**, Woynaroski T, Baranek GT, Wallace MT (2016) Toward an interdisciplinary approach to understanding sensory function in autism spectrum disorder. *Autism Res* **9**:920-925 <https://doi.org/10.1002/aur.1612> | PubMed
- Cerliani L**, Mennes M, Thomas RM, Di Martino A, Thioux M, Keyzers C (2015) Increased functional connectivity between subcortical and cortical resting-state networks in autism spectrum disorder. *JAMA psychiatry* **72**:767-777 <https://doi.org/10.1001/jamapsychiatry.2015.0101> | PubMed
- Chamak B**, Bonniau B, Jaunay E, Cohen D (2008) What can we learn about autism from autistic persons?. *Psychotherapy and psychosomatics* **77**:271-279 <https://doi.org/10.1159/000140086> | PubMed
- Chen YJ**, Duku E, Georgiades S (2022) Rethinking autism intervention science: A dynamic perspective. *Frontiers in psychiatry* **13**:827406 <https://doi.org/10.3389/fpsy.2022.827406> | PubMed
- Cohen SS**, Tottenham N, Baldassano C (2022) Developmental changes in story-evoked responses in the neocortex and hippocampus. *eLife* **11** <https://doi.org/10.7554/elife.69430> | PubMed
- Colavita FB** (1974) Human sensory dominance. *Perception & Psychophysics* **16**:409-412 <https://doi.org/10.3758/bf03203962>
- Constantino JN** (2007) *Social Responsiveness Scale (SRS)* Western Psychological Services.

- Constantino JN**, Davis SA, Todd RD, Schindler MK, Gross MM, Brophy SL, Metzger LM, Shoushtari CS, Splinter R, Reich W (2003) Validation of a brief quantitative measure of autistic traits: comparison of the social responsiveness scale with the autism diagnostic interview-revised. *J Autism Dev Disord* **33**:427-433 <https://doi.org/10.1023/a:1025014929212> | PubMed
- Courchesne E**, Pierce K (2005) Why the frontal cortex in autism might be talking only to itself: local over-connectivity but long-distance disconnection. *Curr Opin Neurobiol* **15**:225-230 <https://doi.org/10.1016/j.conb.2005.03.001> | PubMed
- Van de Cruys S**, Evers K, Van der Hallen R, Van Eylen L, Boets B, de Wit L, Wagemans J (2014) Precise minds in uncertain worlds: predictive coding in autism. *Psychol Rev* **121**:649-675 <https://doi.org/10.1037/a0037665> | PubMed
- Dakin S**, Frith U (2005) Vagaries of visual perception in autism. *Neuron* **48**:497-507 <https://doi.org/10.1016/j.neuron.2005.10.018> | PubMed
- Damarla SR**, Keller TA, Kana RK, Cherkassky VL, Williams DL, Minshew NJ, Just MA (2010) Cortical underconnectivity coupled with preserved visuospatial cognition in autism: Evidence from an fMRI study of an embedded figures task. *Autism Res* **3**:273-279 <https://doi.org/10.1002/aur.153> | PubMed
- Di X**, Xu T, Uddin LQ, Biswal BB (2023) Individual differences in time-varying and stationary brain connectivity during movie watching from childhood to early adulthood: Age, sex, and behavioral associations. *Dev Cogn Neurosci* **63**:101280 <https://doi.org/10.1016/j.dcn.2023.101280> | PubMed
- Di Martino A**, Yan CG, Li Q, Denio E, Castellanos FX, Alaerts K, Anderson JS, Assaf M, Bookheimer SY, Dapretto M, et al. (2014) The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. *Mol Psychiatry* **19**:659-667 <https://doi.org/10.1038/mp.2013.78> | PubMed
- Diedrichsen J**, Kriegeskorte N (2017) Representational models: A common framework for understanding encoding, pattern-component, and representational-similarity analysis. *PLoS Comput Biol* **13**:e1005508 <https://doi.org/10.1371/journal.pcbi.1005508> | PubMed
- Dinstein I**, Heeger DJ, Lorenzi L, Minshew NJ, Malach R, Behrmann M (2012) Unreliable evoked responses in autism. *Neuron* **75**:981-991 <https://doi.org/10.1016/j.neuron.2012.07.026> | PubMed
- DuBois D**, Ameis SH, Lai MC, Casanova MF, Desarkar P (2016) Interoception in autism spectrum disorder: A review. *International journal of developmental neuroscience* **52**:104-111 <https://doi.org/10.1016/j.ijdevneu.2016.05.001> | PubMed
- Elyounssi S**, Kunitoki K, Clauss JA, Laurent E, Kane KA, Hughes DE, Hopkinson CE, Bazer O, Sussman RF, Doyle AE, et al. (2025) Addressing artifactual bias in large, automated MRI analyses of brain development. *Nature Neuroscience* 1-10 <https://doi.org/10.1038/s41593-025-01990-7> | PubMed
- Esteban O**, Markiewicz CJ, Blair RW, Moodie CA, Isik AI, Erramuzpe A, Kent JD, Goncalves M, DuPre E, Snyder M, et al. (2019) fMRIPrep: a robust preprocessing pipeline for functional MRI. *Nat Methods* **16**:111-116 <https://doi.org/10.1038/s41592-018-0235-4> | PubMed
- Falck-Ytter T**, Bussu G (2023) The sensory-first account of autism. *Neurosci Biobehav Rev* **153**:105405 <https://doi.org/10.1016/j.neubiorev.2023.105405> | PubMed
- Federici A**, Parma V, Vicovaro M, Radassao L, Casartelli L, Ronconi L (2020) Anomalous perception of biological motion in autism: a conceptual review and meta-analysis. *Scientific Reports* **10**:4576 <https://doi.org/10.1038/s41598-020-61252-3> | PubMed
- Feldman JI**, Conrad JG, Kuang W, Tu A, Liu Y, Simon DM, Wallace MT, Woynaroski TG (2022) Relations between the McGurk effect, social and communication skill, and autistic features in children with and without autism. *J Autism Dev Disord* **52**:1920-1928 <https://doi.org/10.1007/s10803-021-05074-w> | PubMed
- Feldman JI**, Dunham K, Cassidy M, Wallace MT, Liu Y, Woynaroski TG (2018) Audiovisual multisensory integration in individuals with autism spectrum disorder: A systematic review and meta-analysis. *Neurosci Biobehav Rev* **95**:220-234 <https://doi.org/10.1016/j.neubiorev.2018.09.020> | PubMed

- Felleman DJ, Van Essen DC (1991) Distributed hierarchical processing in the primate cerebral cortex. *Cereb Cortex* **1**:1-47 <https://doi.org/10.1093/cercor/1.1.1-a> | PubMed
- Floris DL, Llera A, Zabihi M, Moessnang C, Jones EJ, Mason L, Haartsen R, Holz NE, Mei T, Elleaume C, *et al.* (2025) A multimodal neural signature of face processing in autism within the fusiform gyrus. *Nature Mental Health* **3**:31-45 <https://doi.org/10.1038/s44220-024-00349-4> | PubMed
- Friston K (2005) A theory of cortical responses. *Philosophical transactions of the Royal Society B: Biological sciences* **360**:815-836 <https://doi.org/10.1098/rstb.2005.1622> | PubMed
- Frith U (1989) *Autism: Explaining the Enigma* Oxford, UK: Blackwell.
- Gemmeke JF, Ellis DPW, Freedman D, Jansen A, Lawrence W, Moore RC, Plakal M, Ritter M. (2017) Audio Set: An ontology and human-labeled dataset for audio events. In: *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). pp. 776-780 <https://doi.org/10.1109/icassp.2017.7952261>
- Georgiades S, Bishop SL, Frazier T (2017) *Editorial Perspective: Longitudinal research in autism-introducing the concept of 'chronogeneity'* Wiley Online Library.
- Ghazanfar AA, Schroeder CE (2006) Is neocortex essentially multisensory?. *Trends Cogn Sci* **10**:278-285 <https://doi.org/10.1016/j.tics.2006.04.008> | PubMed
- Glasser MF, Coalson TS, Robinson EC, Hacker CD, Harwell J, Yacoub E, Ugurbil K, Andersson J, Beckmann CF, Jenkinson M, *et al.* (2016) A multi-modal parcellation of human cerebral cortex. *Nature* **536**:171-178 <https://doi.org/10.1038/nature18933> | PubMed
- Guo X, He C, Duan X, Han S, Xiao J, Chen H. (2019) Aberrant functional connectivity dynamics of superior temporal sulcus and its associations with GABA genes expression in autism. In: *Proceedings of the 3rd International Conference on Medical and Health Informatics*. pp. 21-25 <https://doi.org/10.1145/3340037.3340049>
- Hadjikhani N, Joseph RM, Snyder J, Tager-Flusberg H (2007) Abnormal activation of the social brain during face perception in autism. *Hum Brain Mapp* **28**:441-449 <https://doi.org/10.1002/hbm.20283> | PubMed
- Han S, Tai C, Westenbroek RE, Yu FH, Cheah CS, Potter GB, Rubenstein JL, Scheuer T, De La Iglesia HO, Catterall WA (2012) Autistic-like behaviour in Scn1a<sup>+/-</sup> mice and rescue by enhanced GABA-mediated neurotransmission. *Nature* **489**:385-390 <https://doi.org/10.1038/nature11356> | PubMed
- Happé FG (1996) Studying weak central coherence at low levels: children with autism do not succumb to visual illusions. A research note. *Journal of child psychology and psychiatry* **37**:873-877 <https://doi.org/10.1111/j.1469-7610.1996.tb01483.x> | PubMed
- Happé F, Ronald A, Plomin R (2006) Time to give up on a single explanation for autism. *Nat Neurosci* **9**:1218-1220 <https://doi.org/10.1038/nn1770> | PubMed
- Hasson U, Avidan G, Gelbard H, Vallines I, Harel M, Minshew N, Behrmann M (2009) Shared and idiosyncratic cortical activation patterns in autism revealed under continuous real-life viewing conditions. *Autism Res* **2**:220-231 <https://doi.org/10.1002/aur.89> | PubMed
- Haxby JV, Connolly AC, Guntupalli JS (2014) Decoding neural representational spaces using multivariate pattern analysis. *Annu Rev Neurosci* **37**:435-456 <https://doi.org/10.1146/annurev-neuro-062012-170325> | PubMed
- Haxby JV, Guntupalli JS, Connolly AC, Halchenko YO, Conroy BR, Gobbini MI, Hanke M, Ramadge PJ (2011) A common, high-dimensional model of the representational space in human ventral temporal cortex. *Neuron* **72**:404-416 <https://doi.org/10.1016/j.neuron.2011.08.026> | PubMed
- Herbert MR, Harris GJ, Adrien KT, Ziegler DA, Makris N, Kennedy DN, Lange NT, Chabris CF, Bakardjiev A, Hodgson J, *et al.* (2002) Abnormal asymmetry in language association cortex in autism. *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society* **52**:588-596 <https://doi.org/10.1002/ana.10349> | PubMed

- Hirst RJ, Cragg L, Allen HA (2018) Vision dominates audition in adults but not children: A meta-analysis of the Colavita effect. *Neurosci Biobehav Rev* **94**:286-301 <https://doi.org/10.1016/j.neubiorev.2018.07.012> | PubMed
- Hollestein V, Poelmans G, Forde NJ, Beckmann CF, Ecker C, Mann C, Schäfer T, Moessnang C, Baumeister S, Banaschewski T, *et al.* (2023) Excitatory/inhibitory imbalance in autism: the role of glutamate and GABA gene-sets in symptoms and cortical brain structure. *Transl Psychiatry* **13**:18 <https://doi.org/10.1038/s41398-023-02317-5> | PubMed
- Huth AG, de Heer WA, Griffiths TL, Theunissen FE, Gallant JL (2016) Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature* **532**:453-458 <https://doi.org/10.1038/nature17637> | PubMed
- Huth AG, Nishimoto S, Vu AT, Gallant JL (2012) A continuous semantic space describes the representation of thousands of object and action categories across the human brain. *Neuron* **76**:1210-1224 <https://doi.org/10.1016/j.neuron.2012.10.014> | PubMed
- Igelström KM, Webb TW, Graziano MS (2017) Functional connectivity between the temporoparietal cortex and cerebellum in autism spectrum disorder. *Cerebral Cortex* **27**:2617-2627 <https://doi.org/10.1093/cercor/bhw079> | PubMed
- Jolly E (2018) Pymer4: Connecting R and python for linear mixed modeling. *J Open Source Softw* **3**:862 <https://doi.org/10.21105/joss.00862>
- Kember J, Patenaude P, Sweatman H, Van Schaik L, Tabuenca Z, Chai XJ (2024) Specialization of anterior and posterior hippocampal functional connectivity differs in autism. *Autism Res* <https://doi.org/10.1002/aur.3170> | PubMed
- Keown CL, Shih P, Nair A, Peterson N, Mulvey ME, Müller RA (2013) Local functional overconnectivity in posterior brain regions is associated with symptom severity in autism spectrum disorders. *Cell reports* **5**:567-572 <https://doi.org/10.1016/j.celrep.2013.10.003> | PubMed
- Khosla M, Ngo GH, Jamison K, Kuceyeski A, Sabuncu MR (2021) Cortical response to naturalistic stimuli is largely predictable with deep neural networks. *Sci Adv* **7** <https://doi.org/10.1126/sciadv.abe7547> | PubMed
- Klin A, Jones W, Schultz R, Volkmar F, Cohen D (2002) Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism. *Archives of general psychiatry* **59**:809-816 <https://doi.org/10.1001/archpsyc.59.9.809> | PubMed
- Koldewyn K, Whitney D, Rivera SM (2011) Neural correlates of coherent and biological motion perception in autism. *Dev Sci* **14**:1075-1088 <https://doi.org/10.1111/j.1467-7687.2011.01058.x> | PubMed
- Kriegeskorte N, Douglas PK (2019) Interpreting encoding and decoding models. *Curr Opin Neurobiol* **55**:167-179 <https://doi.org/10.1016/j.conb.2019.04.002> | PubMed
- Lahnakoski JM, Glerean E, Salmi J, Jääskeläinen IP, Sams M, Hari R, Nummenmaa L (2012) Naturalistic fMRI mapping reveals superior temporal sulcus as the hub for the distributed brain network for social perception. *Front Hum Neurosci* **6**:233 <https://doi.org/10.3389/fnhum.2012.00233> | PubMed
- Lake EMR, Finn ES, Noble SM, Vanderwal T, Shen X, Rosenberg MD, Spann MN, Chun MM, Scheinost D, Constable RT (2019) The Functional Brain Organization of an Individual Allows Prediction of Measures of Social Abilities Transdiagnostically in Autism and Attention-Deficit/Hyperactivity Disorder. *Biol Psychiatry* **86**:315-326 <https://doi.org/10.1016/j.biopsych.2019.02.019> | PubMed
- Lawrence KE, Hernandez LM, Bookheimer SY, Dapretto M (2019) Atypical longitudinal development of functional connectivity in adolescents with autism spectrum disorder. *Autism Research* **12**:53-65 <https://doi.org/10.1002/aur.1971> | PubMed
- Lawson RP, Mathys C, Rees G (2017) Adults with autism overestimate the volatility of the sensory environment. *Nature neuroscience* **20**:1293-1299 <https://doi.org/10.1038/nn.4615> | PubMed
- Lawson RP, Rees G, Friston KJ (2014) An aberrant precision account of autism. *Front Hum Neurosci* **8**:302 <https://doi.org/10.3389/fnhum.2014.00302> | PubMed

- Leekam SR, Nieto C, Libby SJ, Wing L, Gould J (2007) Describing the sensory abnormalities of children and adults with autism. *J Autism Dev Disord* **37**:894-910 <https://doi.org/10.1007/s10803-006-0218-7> | PubMed
- Lefebvre A, Tillmann J, Cliquet F, Amsellem F, Maruani A, Leblond C, Beggiato A, Germanaud D, Amestoy A, Ly-Le Moal M, et al. (2023) Tackling hypo and hyper sensory processing heterogeneity in autism: From clinical stratification to genetic pathways. *Autism Research* **16**:364-378 <https://doi.org/10.1002/aur.2861> | PubMed
- Levit-Binnun N, Davidovitch M, Golland Y (2013) Sensory and motor secondary symptoms as indicators of brain vulnerability. *Journal of Neurodevelopmental Disorders* **5**:26 <https://doi.org/10.1186/1866-1955-5-26> | PubMed
- Lin F, Albantakis L, Noppari T, Santavirta S, Brandi ML, Sun L, Lukkarinen L, Tani P, Salmi J, Nummenmaa L, et al. (2025) Reduced inter-subject functional connectivity during movies in autism: Replicability across cross-national fMRI datasets. *bioRxiv* <https://doi.org/10.1101/2025.02.04.636405>
- Lin R, Naselaris T, Kay K, Wehbe L (2024) Stacked regressions and structured variance partitioning for interpretable brain maps. *Neuroimage* **298**:120772 <https://doi.org/10.1016/j.neuroimage.2024.120772> | PubMed
- Lyons KM, Stevenson RA, Owen AM, Stojanoski B (2020) Examining the relationship between measures of autistic traits and neural synchrony during movies in children with and without autism. *Neuroimage Clin* **28**:102477 <https://doi.org/10.1016/j.nicl.2020.102477> | PubMed
- Marco EJ, Hinkley LBN, Hill SS, Nagarajan SS (2011) Sensory processing in autism: a review of neurophysiologic findings. *Pediatr Res* **69**:48R-54R <https://doi.org/10.1203/pdr.0b013e3182130c54> | PubMed
- Martínez K, Martínez-García M, Marcos-Vidal L, Janssen J, Castellanos FX, Pretus C, Villarroja O, Pina-Camacho L, Díaz-Caneja CM, Parellada M, et al. (2020) Sensory-to-cognitive systems integration is associated with clinical severity in autism spectrum disorder. *J Am Acad Child Adolesc Psychiatry* **59**:422-433 <https://doi.org/10.1016/j.jaac.2019.05.033> | PubMed
- McClure P, Rho N, Lee JA, Kaczmarzyk JR, Zheng CY, Ghosh SS, Nielson DM, Thomas AG, Bandettini P, Pereira F (2019) Knowing what you know in brain segmentation using Bayesian deep neural networks. *Front Neuroinform* **13**:67 <https://doi.org/10.3389/fninf.2019.00067> | PubMed
- McDermott J (2018) pycochleagram: Generate cochleagrams natively in Python. GitHub. <https://github.com/mcdermottLab/pycochleagram?tab=readme-ov-file>
- Mehta K, Salo T, Madison TJ, Adebimpe A, Bassett DS, Bertolero M, Cieslak M, Covitz S, Houghton A, Keller AS, et al. (2024) XCP-D: A robust pipeline for the post-processing of fMRI data. *Imaging Neuroscience* **2**:1-26 [https://doi.org/10.1162/imag\\_a\\_00257](https://doi.org/10.1162/imag_a_00257) | PubMed
- Mihailov A, Philippe C, Gloaguen A, Grigis A, Laidi C, Piquet C, Houenou J, Frouin V (2020) Cortical signatures in behaviorally clustered autistic traits subgroups: a population-based study. *Transl Psychiatry* **10**:207 <https://doi.org/10.1038/s41398-020-00894-3> | PubMed
- Moro SS, Ghemraoui AA, Steeves JKE (2012) No Colavita effect: Lack of visual dominance in people with autism spectrum disorder. *J Vis* **12**:1038-1038 <https://doi.org/10.1167/12.9.1038>
- Mottron L, Burack JA (2001) Enhanced perceptual functioning in the development of autism. In: Burack JA, Charman T, Yirmiya N, Zelazo PR (Eds). *The Development of Autism: Perspectives from Theory and Research* Lawrence Erlbaum Associates Publishers. pp. 131-148
- Mottron L, Dawson M, Soulières I, Hubert B, Burack J (2006) Enhanced perceptual functioning in autism: an update, and eight principles of autistic perception. *J Autism Dev Disord* **36**:27-43 <https://doi.org/10.1007/s10803-005-0040-7> | PubMed
- Mumford D (1992) On the computational architecture of the neocortex: II The role of cortico-cortical loops. *Biological cybernetics* **66**:241-251 <https://doi.org/10.1007/bf00198477> | PubMed

- Napolitano A**, Schiavi S, La Rosa P, Rossi-Espagnet MC, Petrillo S, Bottino F, Tagliente E, Longo D, Lupi E, Casula L, *et al.* (2022) Sex differences in autism spectrum disorder: diagnostic, neurobiological, and behavioral features. *Frontiers in psychiatry* **13**:889636 <https://doi.org/10.3389/fpsy.2022.889636> | [PubMed](#)
- Naselaris T**, Kay KN, Nishimoto S, Gallant JL (2011) Encoding and decoding in fMRI. *Neuroimage* **56**:400-410 <https://doi.org/10.1016/j.neuroimage.2010.07.073> | [PubMed](#)
- Nastase SA**, Goldstein A, Hasson U (2020) Keep it real: rethinking the primacy of experimental control in cognitive neuroscience. *Neuroimage* **222**:117254 <https://doi.org/10.1016/j.neuroimage.2020.117254> | [PubMed](#)
- Nebel MB**, Lidstone DE, Wang L, Benkeser D, Mostofsky SH, Risk BB (2022) Accounting for motion in resting-state fMRI: What part of the spectrum are we characterizing in autism spectrum disorder?. *Neuroimage* **257**:119296 <https://doi.org/10.1016/j.neuroimage.2022.119296> | [PubMed](#)
- Neil PA**, Chee-Ruiter C, Scheier C, Lewkowicz DJ, Shimojo S (2006) Development of multisensory spatial integration and perception in humans. *Developmental science* **9**:454-464 <https://doi.org/10.1111/j.1467-7687.2006.00512.x> | [PubMed](#)
- Nieto Del Rincón PL.** (2008) Autism: alterations in auditory perception. *Rev Neurosci* **19**:61-78 <https://doi.org/10.1515/revneuro.2008.19.1.61> | [PubMed](#)
- Noel JP**, Angelaki DE (2023) A theory of autism bridging across levels of description. *Trends in Cognitive Sciences* **27**:631-641 <https://doi.org/10.1016/j.tics.2023.04.010> | [PubMed](#)
- Nunez-Elizalde A**, la Tour TD, di Oleggio Castello MV (2022) gallantlab/pymoten: v0.0.4. Zenodo.
- O'Connor N**, Hermelin B (1965) Sensory dominance in autistic imbecile children and controls. *Arch Gen Psychiatry* **12**:99-103 <https://doi.org/10.1001/archpsyc.1965.01720310101015>
- Okada K**, Venezia JH, Matchin W, Saberi K, Hickok G (2013) An fMRI Study of Audiovisual Speech Perception Reveals Multisensory Interactions in Auditory Cortex. *PLoS One* **8**:e68959 <https://doi.org/10.1371/journal.pone.0068959> | [PubMed](#)
- O'riordan MA** (2004) Superior visual search in adults with autism. *Autism* **8**:229-248 <https://doi.org/10.1177/1362361304045219> | [PubMed](#)
- Pantelis PC**, Byrge L, Tyszka JM, Adolphs R, Kennedy DP (2015) A specific hypoactivation of right temporo-parietal junction/posterior superior temporal sulcus in response to socially awkward situations in autism. *Soc Cogn Affect Neurosci* **10**:1348-1356 <https://doi.org/10.1093/scan/nsv021> | [PubMed](#)
- Pellicano E**, Burr D (2012) When the world becomes 'too real': a Bayesian explanation of autistic perception. *Trends Cogn Sci* **16**:504-510 <https://doi.org/10.1016/j.tics.2012.08.009> | [PubMed](#)
- Pierce K**, Haist F, Sedaghat F, Courchesne E (2004) The brain response to personally familiar faces in autism: findings of fusiform activity and beyond. *Brain* **127**:2703-2716 <https://doi.org/10.1093/brain/awh289> | [PubMed](#)
- Pierce K**, Müller RA, Ambrose J, Allen G, Courchesne E (2001) Face processing occurs outside the fusiformface area in autism: evidence from functional MRI. *Brain* **124**:2059-2073 <https://doi.org/10.1093/brain/124.10.2059> | [PubMed](#)
- Puts NA**, Wodka EL, Harris AD, Crocetti D, Tommerdahl M, Mostofsky SH, Edden RA (2017) Reduced GABA and altered somatosensory function in children with autism spectrum disorder. *Autism Research* **10**:608-619 <https://doi.org/10.1002/aur.1691> | [PubMed](#)
- Rao RP**, Ballard DH (1999) Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature neuroscience* **2**:79-87 <https://doi.org/10.1038/4580> | [PubMed](#)
- Redcay E**, Moraczewski D (2020) Social cognition in context: A naturalistic imaging approach. *Neuroimage* **216**:116392 <https://doi.org/10.1016/j.neuroimage.2019.116392> | [PubMed](#)
- Robertson CE**, Baron-Cohen S (2017) Sensory perception in autism. *Nat Rev Neurosci* **18**:671-684 <https://doi.org/10.1038/nrn.2017.112> | [PubMed](#)

- Robertson CE, Ratai EM, Kanwisher N (2016) Reduced GABAergic action in the autistic brain. *Current Biology* **26**:80-85 <https://doi.org/10.1016/j.cub.2015.11.019> | PubMed
- Robertson CE, Thomas C, Kravitz DJ, Wallace GL, Baron-Cohen S, Martin A, Baker CI (2014) Global motion perception deficits in autism are reflected as early as primary visual cortex. *Brain* **137**:2588-2599 <https://doi.org/10.1093/brain/awu189> | PubMed
- Ronconi L, Vitale A, Federici A, Mazzoni N, Battaglini L, Molteni M, Casartelli L (2023) Neural dynamics driving audiovisual integration in autism. *Cereb Cortex* **33**:543-556 <https://doi.org/10.1093/cercor/bhac083> | PubMed
- Ross P, Atkins B, Allison L, Simpson H, Duffell C, Williams M, Ermolina O (2021) Children cannot ignore what they hear: Incongruent emotional information leads to an auditory dominance in children. *J Exp Child Psychol* **204**:105068 <https://doi.org/10.1016/j.jecp.2020.105068> | PubMed
- Rubenstein JL, Merzenich MM (2003) Model of autism: increased ratio of excitation/inhibition in key neural systems. *Genes, Brain and Behavior* **2**:255-267 <https://doi.org/10.1034/j.1601-183x.2003.00037.x> | PubMed
- Rudie JD, Dapretto M (2013) Convergent evidence of brain overconnectivity in children with autism?. *Cell reports* **5**:565-566 <https://doi.org/10.1016/j.celrep.2013.10.043> | PubMed
- Russo N, Cascio CJ, Baranek GT, Woynaroski TG, Williams ZJ, Green SA, Schaaf R (2025) A cascading effects model of early sensory development in autism. *Psychological Review* <https://doi.org/10.1037/rev0000558> | PubMed
- Rutter M, Bailey A, Lord C (2003) *SCQ: The Social Communication Questionnaire* Western psychological services.
- Salmi J, Roine U, Glerean E, Lahnakoski J, Nieminen-von Wendt T, Tani P, Leppämäki S, Nummenmaa L, Jääskeläinen IP, Carlson S, et al. (2013) The brains of high functioning autistic individuals do not synchronize with those of others. *Neuroimage Clin* **3**:489-497 <https://doi.org/10.1016/j.nicl.2013.10.011> | PubMed
- Samara A, Zada Z, Vanderwal T, Hasson U, Nastase SA (2025) Cortical language areas are coupled via a soft hierarchy of model-based linguistic features. *bioRxiv* <https://doi.org/10.1101/2025.06.02.657491> | PubMed
- Schneebeli M, Haker H, Rüesch A, Zahnd N, Marino S (2022) Disentangling “Bayesian brain” theories of autism spectrum disorder. *medRxiv* <https://doi.org/10.1101/2022.02.07.22270242>
- Schultz RT, Grelotti DJ, Klin A, Kleinman J, Van der Gaag C, Marois R, Skudlarski P (2003) The role of the fusiform face area in social cognition: implications for the pathobiology of autism. *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences* **358**:415-427 <https://doi.org/10.1098/rstb.2002.1208> | PubMed
- Schwarzkopf DS, Anderson EJ, de Haas B, White SJ, Rees G (2014) Larger extrastriate population receptive fields in autism spectrum disorders. *J Neurosci* **34**:2713-2724 <https://doi.org/10.1523/jneurosci.4416-13.2014> | PubMed
- Shih P, Keehn B, Oram JK, Leyden KM, Keown CL, Müller RA (2011) Functional differentiation of posterior superior temporal sulcus in autism: a functional connectivity magnetic resonance imaging study. *Biological psychiatry* **70**:270-277 <https://doi.org/10.1016/j.biopsych.2011.03.040> | PubMed
- Simonoff E, Kent R, Stringer D, Lord C, Briskman J, Lukito S, Pickles A, Charman T, Baird G (2020) Trajectories in symptoms of autism and cognitive ability in autism from childhood to adult life: Findings from a longitudinal epidemiological cohort. *Journal of the American Academy of Child & Adolescent Psychiatry* **59**:1342-1352 <https://doi.org/10.1016/j.jaac.2019.11.020> | PubMed
- Smith LE, Maenner MJ, Seltzer MM (2012) Developmental trajectories in adolescents and adults with autism: The case of daily living skills. *Journal of the American Academy of Child & Adolescent Psychiatry* **51**:622-631 <https://doi.org/10.1016/j.jaac.2012.03.001> | PubMed

- Sokolov AA**, Erb M, Gharabaghi A, Grodd W, Tatagiba MS, Pavlova MA (2012) Biological motion processing: the left cerebellum communicates with the right superior temporal sulcus. *Neuroimage* **59**:2824-2830 <https://doi.org/10.1016/j.neuroimage.2011.08.039> | PubMed
- Sonkusare S**, Breakspear M, Guo C (2019) Naturalistic Stimuli in Neuroscience: Critically Acclaimed. *Trends Cogn Sci* **23**:699-714 <https://doi.org/10.1016/j.tics.2019.05.004> | PubMed
- Stein BE**, Stanford TR (2008) Multisensory integration: current issues from the perspective of the single neuron. *Nature reviews neuroscience* **9**:255-266 <https://doi.org/10.1038/nrn2331> | PubMed
- Steinmetz CJ** (2023) pyloudnorm: Flexible audio loudness meter in Python with implementation of ITU-R BS.1770-4 loudness algorithm.
- Stevenson RA**, Segers M, Ferber S, Barense MD, Wallace MT (2014) The impact of multisensory integration deficits on speech perception in children with autism spectrum disorders. *Front Psychol* **5**:379 <https://doi.org/10.3389/fpsyg.2014.00379> | PubMed
- Stevenson RA**, Siemann JK, Woynarowski TG, Schneider BC, Eberly HE, Camarata SM, Wallace MT (2014) Evidence for diminished multisensory integration in autism spectrum disorders. *J Autism Dev Disord* **44**:3161-3167 <https://doi.org/10.1007/s10803-014-2179-6> | PubMed
- Supekar K**, Uddin LQ, Khouzam A, Phillips J, Gaillard WD, Kenworthy LE, Yerys BE, Vaidya CJ, Menon V (2013) Brain hyperconnectivity in children with autism and its links to social deficits. *Cell Rep* **5**:738-747 <https://doi.org/10.1016/j.celrep.2013.10.001> | PubMed
- Sydnor VJ**, Larsen B, Bassett DS, Alexander-Bloch A, Fair DA, Liston C, Mackey AP, Milham MP, Pines A, Roalf DR, *et al.* (2021) Neurodevelopment of the association cortices: Patterns, mechanisms, and implications for psychopathology. *Neuron* **109**:2820-2846 <https://doi.org/10.1016/j.neuron.2021.06.016> | PubMed
- Takesian AE**, Hensch TK (2013) Balancing plasticity/stability across brain development. *Progress in brain research* **207**:3-34 <https://doi.org/10.1016/b978-0-444-63327-9.00001-1> | PubMed
- Tavassoli T**, Miller LJ, Schoen SA, Nielsen DM, Baron-Cohen S (2014) Sensory over-responsivity in adults with autism spectrum conditions. *Autism* **18**:428-432 <https://doi.org/10.1177/1362361313477246> | PubMed
- Thompson RA**, Nelson CA (2001) Developmental science and the media: Early brain development. *American Psychologist* **56**:5 <https://doi.org/10.1037/0003-066x.56.1.5> | PubMed
- Tong X**, Xie H, Fonzo GA, Zhao K, Satterthwaite TD, Carlisle NB, Zhang Y (2024) Symptom dimensions of resting-state electroencephalographic functional connectivity in autism. *Nature Mental Health* 1-12 <https://doi.org/10.1038/s44220-023-00195-w> | PubMed
- Dupré la Tour T**, Eickenberg M, Nunez-Elizalde AO, Gallant JL (2022) Feature-space selection with banded ridge regression. *Neuroimage* **264**:119728 <https://doi.org/10.1016/j.neuroimage.2022.119728> | PubMed
- Dupré la Tour T**, Visconti di Oleggio Castello M, Gallant JL (2025) The Voxelwise Encoding Model framework: A tutorial introduction to fitting encoding models to fMRI data. *Imaging Neuroscience* [https://doi.org/10.1162/imag\\_a\\_00575](https://doi.org/10.1162/imag_a_00575) | PubMed
- Uljarević M**, Baranek G, Vivanti G, Hedley D, Hudry K, Lane A (2017) Heterogeneity of sensory features in autism spectrum disorder: Challenges and perspectives for future research. *Autism Research* **10**:703-710 <https://doi.org/10.1002/aur.1747> | PubMed
- Vanderwal T**, Eilbott J, Castellanos FX (2019) Movies in the magnet: Naturalistic paradigms in developmental functional neuroimaging. *Dev Cogn Neurosci* **36**:100600 <https://doi.org/10.1016/j.dcn.2018.10.004> | PubMed
- Wang X**, Bouton S, Kojovic N, Giraud AL, Schaer M (2024) Atypical audio-visual neural synchrony and speech processing in children with autism spectrum disorder. *bioRxiv* 2024.04.19.590044 <https://doi.org/10.1101/2024.04.19.590044>

Wei H, Zhu Y, Wang T, Zhang X, Zhang K, Zhang Z (2021) Genetic risk factors for autism-spectrum disorders: A systematic review based on systematic reviews and meta-analysis. *Journal of Neural Transmission* **128**:717-734 <https://doi.org/10.1007/s00702-021-02360-w> | PubMed

Williams ZJ, Schaaf R, Ausderau KK, Baranek GT, Barrett DJ, Cascio CJ, Dumont RL, Eyoh EE, Failla MD, Feldman JI, et al. (2023) Examining the latent structure and correlates of sensory reactivity in autism: a multi-site integrative data analysis by the autism sensory research consortium. *Mol Autism* **14**:31 <https://doi.org/10.1186/s13229-023-00563-4> | PubMed

Xiao Y, Friederici AD, Margulies DS, Brauer J (2016) Longitudinal changes in resting-state fMRI from age 5 to age 6 years covary with language development. *Neuroimage* **128**:116-124 <https://doi.org/10.1016/j.neuroimage.2015.12.008> | PubMed

You W, Li Q, Chen L, He N, Li Y, Long F, Wang Y, Chen Y, McNamara RK, Sweeney JA, et al. (2024) Common and distinct cortical thickness alterations in youth with autism spectrum disorder and attention-deficit/hyperactivity disorder. *BMC medicine* **22**:92 <https://doi.org/10.1186/s12916-024-03313-2> | PubMed

Lindsay M.Alexander et al (2017) Healthy Brain Network. Loris. <https://data.healthybrainnetwork.org/>

## Peer reviews

### Reviewer #1 (Public review):

#### Summary:

This study uses stacked encoding models to characterize differences in sensory (visual and auditory) processing between autistic and non-autistic children and adolescents. The authors found no significant enhancement of low-level feature encoding in either visual or auditory cortex, but reduced high-level visual representations and a relative shift toward low-level over high-level visual feature encoding in the posterior superior temporal sulcus (pSTS). The shift in pSTS correlated with social symptom severity (SRS scores). These findings support weak central coherence (WCC) theory over enhanced perceptual functioning (EPF) theory, suggesting an altered visual feature encoding in pSTS in autism.

#### Strengths:

This study uses sophisticated methodology and an open data set with a relatively large sample size. fMRI data are acquired during a naturalistic paradigm (i.e., movie watching), which promotes attention and engagement among participants, and provides greater ecological validity. The use of encoding models to explore population-level differences in neural representations of stimulus-computable features is novel. Overall, results provide somewhat modest yet still informative evidence for adjudicating between possible theories of altered sensory processing in autism.

#### Weaknesses:

Some important methodological details are missing and/or require justification. Some potential confounding factors or unconsidered differences between individuals and/or diagnostic groups should be explored and possibly addressed. Specific major and minor points are raised below.

#### Major comments:

(1) Unclear description of noise ceiling calculation (line 205-206, 632-634) and potential heterogeneity: it is not clear what data were "split" for the split-half correlation used to calculate noise ceilings. To our knowledge, each participant watched each movie once each, so there is no within-subject repetition available. Were these correlations across participants

(i.e., ISC)? If so, does this across-subject metric provide a fair representation of the true noise ceiling, given that a) encoding models themselves are trained within subjects and b) autistic individuals are known to exhibit more idiosyncrasy in responses to naturalistic stimuli (e.g., Hasson et al., 2008)? Moreover, do noise ceilings differ between individual participants, diagnostic groups, and/or with age? If so, how might these differences affect the interpretation of results (e.g., R2 differences)?

(2) Possibly underperforming visual model: given that the visual model in general performed worse than the audio model, the visual vs audio perceptual preference analyses (line 281-290) might be affected by the underlying mismatch between model performance. Though the visual and auditory regions showed similar noise ceilings (Figure 2 S1B), the stacked model performed better in auditory regions than in visual or multimodal regions (Figure 2 S1A). Supporting the same idea, the visual model in general showed lower fitting R2 than the audio model (Figure 2 S2A, Figure 2 S3A vs B). Instead of using mean motion (line 608-614), applying PCA on the raw features might help reduce noise inherent in the raw motion energy features (Malik et al., 2026), therefore improving model performance.

(3) The clipping procedure for unique variance (lines 634-637) requires justification: the unique variance is defined by subtracting high-level R<sup>2</sup> from stacked R<sup>2</sup> with explicit clipping when high-level R<sup>2</sup> is negative or exceeds stacked R<sup>2</sup>. However, in the original stacked regression framework (Lin et al., 2024), unique variance is defined by simple subtraction without such post-hoc adjustment, as the negative R<sup>2</sup> is still meaningful, indicating the model performs worse than predicting using the mean value. This requires justification. How frequently does clipping occur, and in which brain regions? Is it an indicator of overfitting or poor model performance? How substantially do results change if clipping is removed? E.g., the hemisphere dominance comparison (line 271-280, Figure 6). Critically, does this procedure affect the key finding regarding SRS/sensory symptom severity correlations in pSTS?

(4) The interpretation of the correlation between SRS with neural patterns is misleading (line 237-242, line 364-366): based on Figure 3, SRS and SSS showed more significant and robust relationship with unique variance of high-level visual feature, meaning that the decrement of high-level feature encoding in STSvp and STSdp, rather than the relative low-level preference, is likely driving the relationship with autism severity and sensory symptom.

(5) Details are missing about how data from the two movie runs were combined. Were the time series concatenated without regard to which movie they originally came from, or was the distinction between movies taken into account for purposes of splitting data into train/test cross-validation folds? The results would be stronger if the authors could show that results replicate across the two movies when they are each analyzed independently, though we recognize that there is perhaps not enough data, especially in the shorter [~4min] movie, to do this. The authors discussed this in lines 412-417, but it would be helpful to provide a justification in the Methods section as well.

(6) Potential feature weight differences across individuals and/or diagnostic categories: since the encoding models were trained for each subject, is there significant variability in feature weights across individuals and/or diagnostic categories (e.g., did the model predictions heavily rely on face for the non-ASD group but not for the ASD group)? If so, how does this change the interpretation of the R2 comparisons? The authors showed the results of stacked feature weight differences between diagnostic categories and their relationship with autism severity and sensory symptoms, but it might be informative to show the raw feature weightings before diving into stacked-weight differences.

<https://doi.org/10.7554/eLife.111008.1.sa2>

## Reviewer #2 (Public review):

### Summary:

This study by Mentch et al. uses naturalistic-movie fMRI and grayordinate-level stacked encoding models to test preregistered hypotheses about low/high-level and audio/visual feature encoding in autism and adolescence from openly available Healthy Brain Network data. Null results reported that autism was not linked to increased low-level encoding in primary sensory cortices. Exploratory analyses showed participants with autism showed reduced high-level visual encoding in social regions (pSTS, face areas), with the high-low feature shift tracking social responsiveness scale (SRS) scores. Age and laterality effects were also found.

### Strengths:

- (1) This study and hypotheses were preregistered.
- (2) The study utilised proper variance partitioning, split-half noise ceilings, FD-threshold sensitivity analyses, and an explicit modelling framework that recovers known sensory hierarchies in the aggregated sample. The developmental sampling adds to the interest.
- (3) The manuscript is written clearly, laying out the background and theories to be tested with encoding models. The analyses and reporting of results are clear.

### Weaknesses:

- (1) If I understand correctly, by only averaging the grayordinates that already passed a significance threshold, the resulting parcel value is guaranteed to look stronger than if all grayordinates had been included. This has been raised in neuroimaging (Kriegeskorte et al., 2009; Vul et al., 2009). Can the authors justify these choices?
- (2) I assume that the phrase "temporally permuting the order of observations" on Page 22 means random shuffling of time points. The details of this exact permutation are not specified. Both the fMRI BOLD signal and movie features have strong temporal autocorrelation, and random shuffling will destroy this structure. This is important as grayordinate-level survivors will propagate to parcel pools. Circular shifting or phase randomization preserving the autocorrelation spectrum is appropriate.
- (3) In the movie feature selection, the low-level visual model contains only two scalars: mean perceptual brightness and a single averaged value across 2,139 motion-energy filters. With only two low-level visual features, the low-level visual model potentially would underestimate low-level visual encoding. The H1.1 toward the null perhaps suggests to this. Principal components of the motion-energy outputs, as was done for the cochleagram, could be used.
- (4) The pilot sample composition is not described. Features were selected based on their performance on an independent set of 54 pilot subjects. Please provide age, sex, and diagnostic composition of the pilot sample. The main point being whether the selected features were optimised for a population that differs from the subject studied.
- (5) The authors acknowledge the lack of eye-tracking in theory study. I think this should be elaborated, especially why this modality is important for answering sensory and perceptual encoding. Face encoding may not be degraded, but just that faces are not being attended to.
- (6) I think a more nuanced distinction about the representational nature of encoding-model  $R^2$  should be mentioned, especially when the interpretation of findings is related to

perceptual functioning (EPF theory).  $R^2$  measures how well a feature set predicts brain activity, not perceptual function or cognitive integration.

(7) The literature also includes evidence for no Colavita effect, not just reverse Colavita in autism, and the framing should reflect this more even-handedly.

(8) The 0.2 mm per-volume threshold is quite strict. The 40%/60%/80% sensitivity analyses partially address this, but a brief justification for the choice of 0.2 mm would strengthen the Methods.

(9) Figure 1 seems confusing and would benefit from more information or text in the figure.

(10) Figure 2 supplement has caption A labelled twice; please correct.

(11) Acronyms. Please spell out MSI on first mention (page 2) and ISC/ISFC on first mention (page 4).

<https://doi.org/10.7554/eLife.111008.1.sa1>

### Reviewer #3 (Public review):

Summary:

This study investigates the neural mechanisms underlying sensory-perceptual differences in autism through a naturalistic movie-viewing fMRI paradigm. By employing encoding models, the authors demonstrate that autistic children and adolescents exhibit a specific alteration in visual feature weighting, characterized by a shift toward low-level visual feature encoding in higher-order association regions, particularly the posterior Superior Temporal Sulcus (pSTS). This shift is linked to social symptom severity, providing empirical support for Weak Central Coherence accounts.

Strengths:

The study's primary strengths lie in its methodological rigor and innovative approach. The use of a pre-registered analysis plan ensures transparency and enhances the credibility of the findings, while the encoding models allow for a fine-grained dissociation of low-level versus high-level feature representations across the cortex. Overall, the writing is clear, the logic is sound, and the results offer a significant contribution to the field by refining our understanding of how sensory processing is differentially organized in autism.

Weaknesses:

While the study presents compelling findings regarding visual feature encoding in autism, several methodological and interpretive limitations warrant consideration. First, the Discussion focuses primarily on WCC and EPF theories, failing to explicitly address how the results intersect with other prominent frameworks mentioned in the Introduction, such as Bayesian predictive coding or E/I imbalance hypotheses. Second, the demographic characteristics and specific sample sizes of the ASD-ADHD and ASD+ADHD subgroups are not reported, limiting the interpretability of the stratified analyses; furthermore, the counterintuitive finding that the ASD+ADHD group resembles controls is not sufficiently discussed. Third, given the significant group difference in IQ and the known relationship between cognitive ability and neural processing, the potential confounding influence of IQ on the neuroimaging results requires more explicit acknowledgment, particularly since IQ was not included as a covariate in the primary models.

<https://doi.org/10.7554/eLife.111008.1.sa0>