Functional neuroimaging identifies a cross-species biomarker for treatment monitoring in Angelman syndrome

Supplementary Methods

fNIRS Data Processing-Method 2

Channel quality assessment was performed using QT-NIRS metrics implemented in MNE-NIRS[1–3]. After raw light-intensity signals were converted to optical density, the Scalp Coupling Index (SCI) was computed from cardiac-band correlations between wavelengths, and channels with SCI < 0.5 were excluded. A Peak Spectral Power metric was additionally computed in the 0.7–1.5 Hz band using 10-s windows, 0.3 Hz transition bandwidths, and a threshold of 0.05; windows falling below threshold were marked as low quality, and overlapping epochs were discarded. Motion correction was implemented using Homer3 functions, which were custom-adapted for compatibility with our Python processing pipeline[4]. Motion artifacts were detected using int_motion = 1.0 s, t_mask = 1.0 s, stdev_thresh = 13.0, and amp_thresh = 0.4. Identified segments were corrected using spline interpolation (p = 0.99), followed by wavelet filtering (Daubechies-2, IQR threshold = 0.8). Optical-density data were converted to concentration changes in HbO and HbR using the modified Beer–Lambert law with a partial pathlength factor (PPF) of 6.0. Signals were then band-pass filtered between 0.01 and 0.20 Hz (transition bandwidth = 0.01 Hz). Correlation-based signal improvement (CBSI) was subsequently applied to enhance the expected HbO–HbR anticorrelation[5]. Following within-participant trial averaging, a top-k channel selection procedure was used as a functional channel selection step to enhance selectivity[6]. Channels were ranked based on the mean HbO response over the 0–20 s post-stimulus interval, and the top 11 channels (50% of the 22-channel montage) were retained. Selected HbO channels were mapped onto the corresponding HbR channels, and total hemoglobin was computed as HbT = HbO + HbR. Retained channels were then averaged within each participant to obtain a single subject-level trace per chromophore and condition.

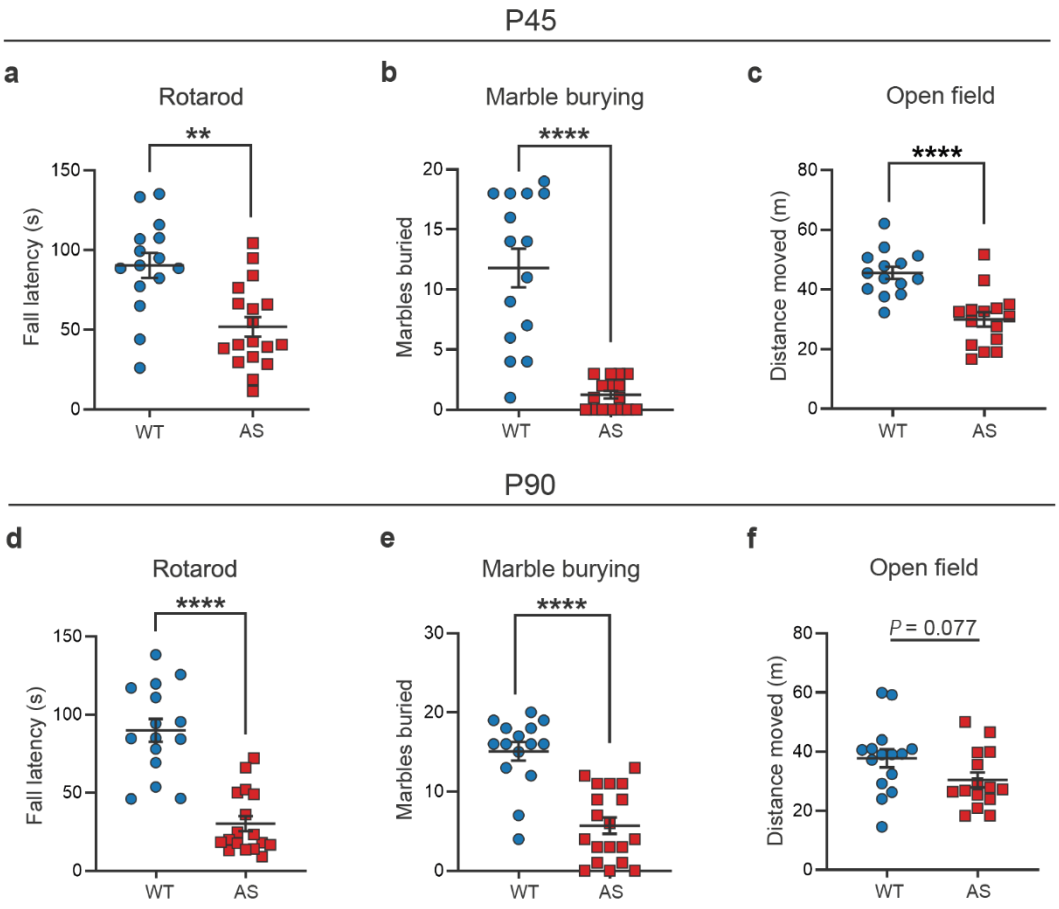
Response Shape and Age-Related Analyses

To quantify the conformity of the hemodynamic response to a canonical profile, HbO stimulation responses from AS participants were compared, for each preprocessing strategy, to a model response obtained by convolving a 5 s boxcar stimulus with a gamma hemodynamic response function. A 3 s delay was applied to account for neurovascular coupling latency, and the resulting waveform was normalized to unit peak. Subject-level responses were baseline-corrected and rescaled to the [0–1] range prior to comparison. Similarity was quantified using root-mean-square error (RMSE), with lower values indicating closer agreement with the canonical response. As both preprocessing strategies were evaluated within the same participants, differences in RMSE were assessed using a one-sided paired t-test.

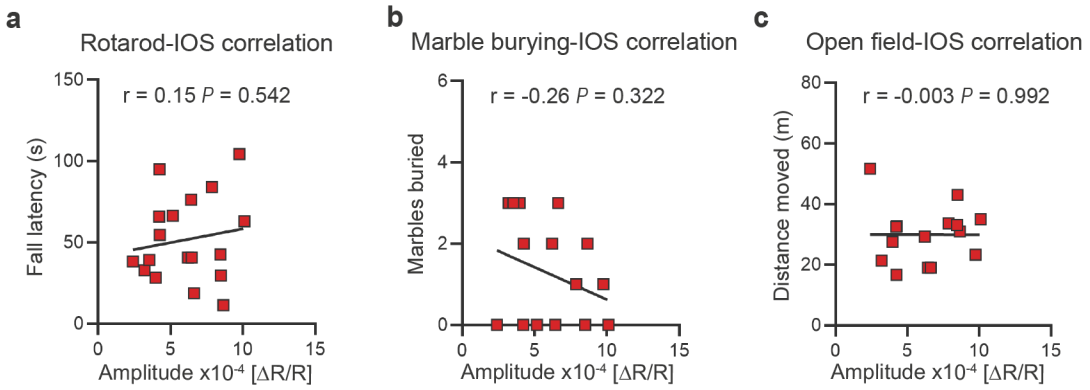
Age-related analyses were restricted to stimulation-derived AUC and peak-amplitude measures and were performed separately within control and Angelman groups using Spearman rank correlation. To estimate the disease-related component of the response while accounting for normative age effects, a smoothing spline was fitted to control data to derive an age-dependent reference curve. Each Angelman participant was then projected onto this model to obtain the age-expected response, and the difference between observed and expected values was defined as the age-normalized excess response. Associations between excess response and age were assessed using linear regression, with one-sided tests evaluating directional hypotheses (slope < 0 and slope > 0). In addition, excess-response values were tested against zero using a one-sample one-sided t-test (mean > 0) to determine whether a positive residual

response persisted after age normalization. When multiple metrics were analyzed, p-values were adjusted using the Benjamini–Hochberg procedure.

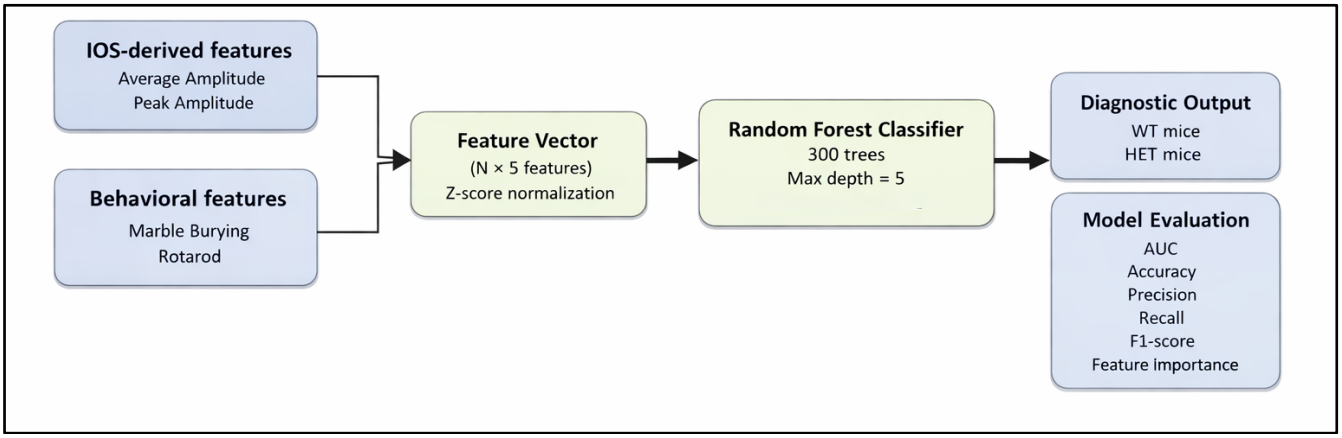
Supplementary Figures



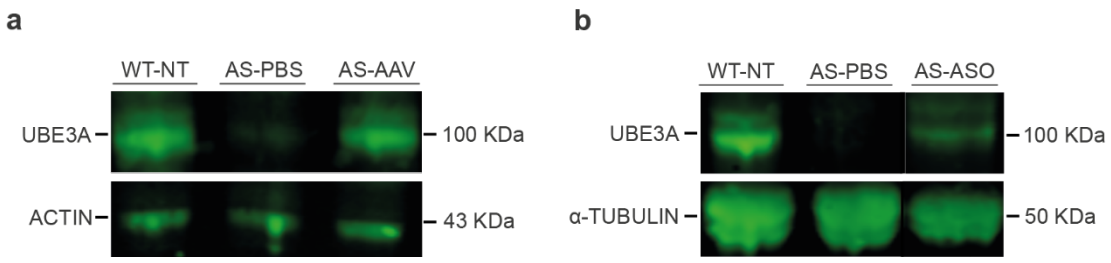
Supplementary Figure 1. Behavioral assessment in juvenile and adult mice. Behavioral assessment of WT and AS juvenile (a-c) and adult mice (d-f) in the rotarod (a, d), marble burying (b, e), and open field (c, f). WT animals are shown as blue circles, while AS mice are shown as red squares. Statistical details are reported in Supplementary Table 1. **p<0.01, ****p< 0.0001



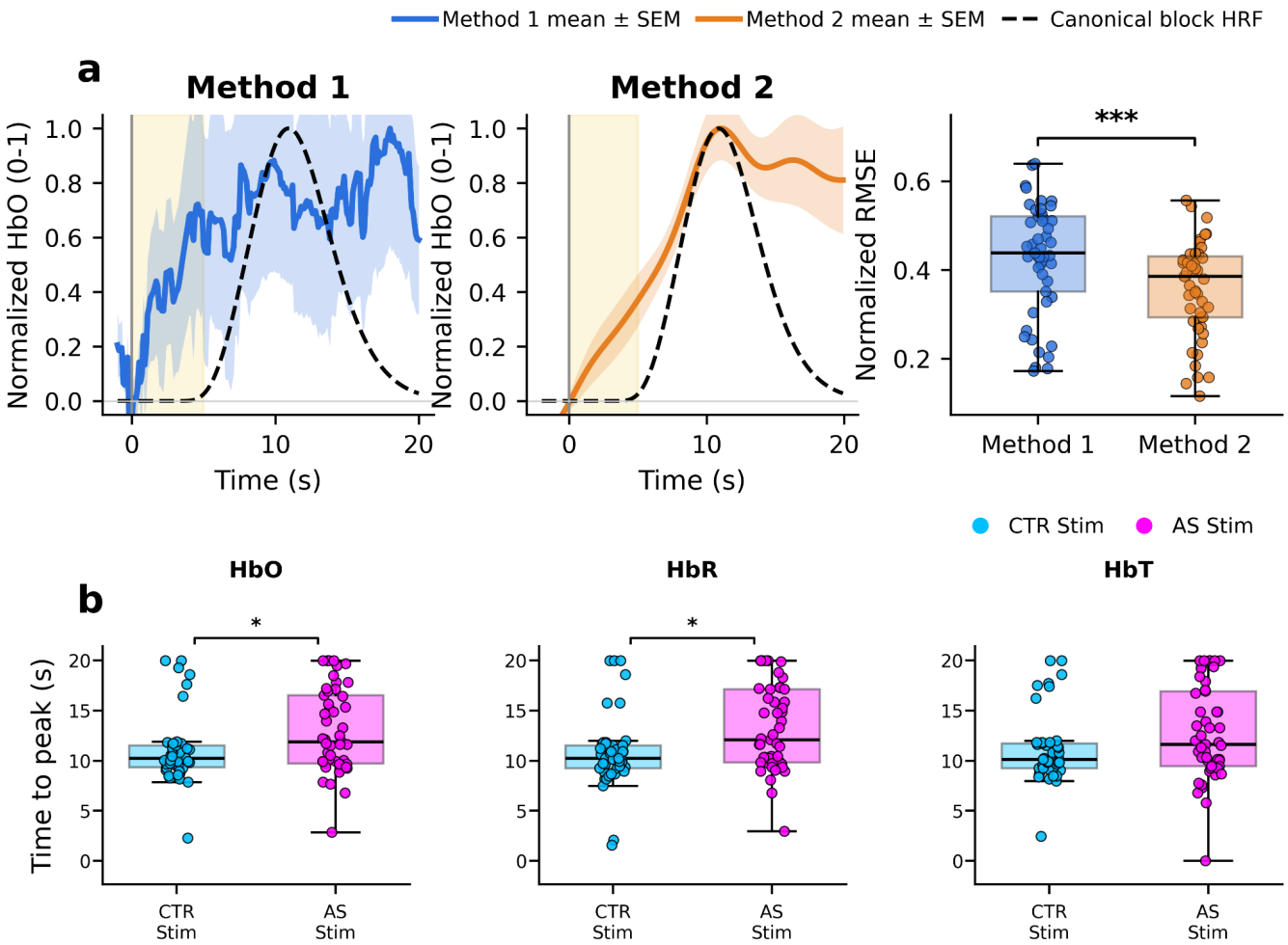
Supplementary Figure 2. Correlation between IOS responses and behavioral performance in juvenile AS mice. Relationship between IOS responses and behavioral performance in the rotarod (a), marble burying (b), and open field (c) in AS juvenile mice. Statistical details are reported in Supplementary Table 1.



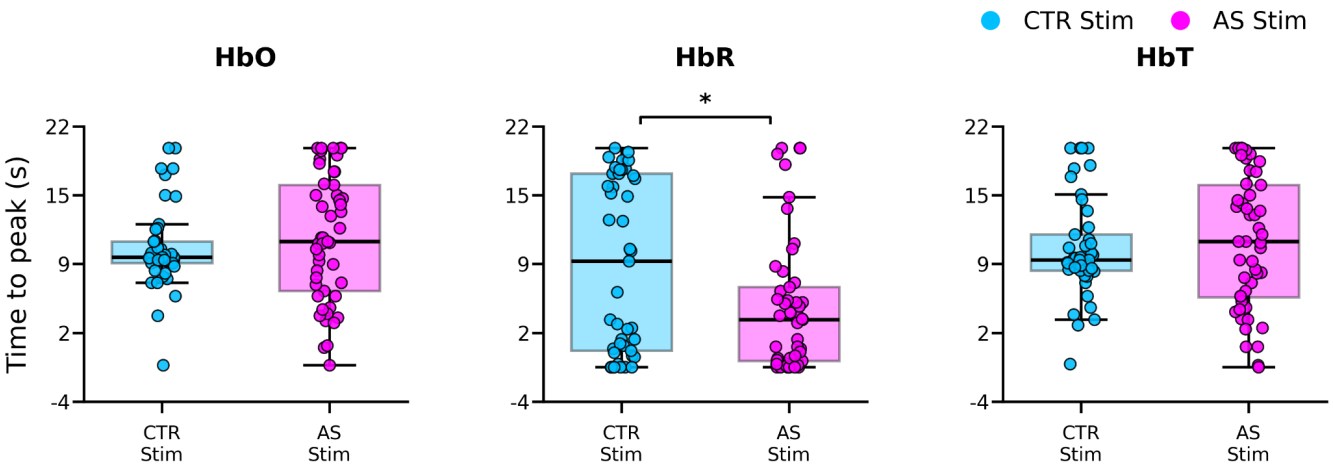
Supplementary Figure 3. Random Forest classification pipeline for WT and AS mice based on IOS and behavioral features. Schematic overview of the machine learning framework illustrating feature extraction from IOS imaging and behavioral assays, followed by feature normalization, random forest classification, and evaluation of model performance used to distinguish WT from AS mice.



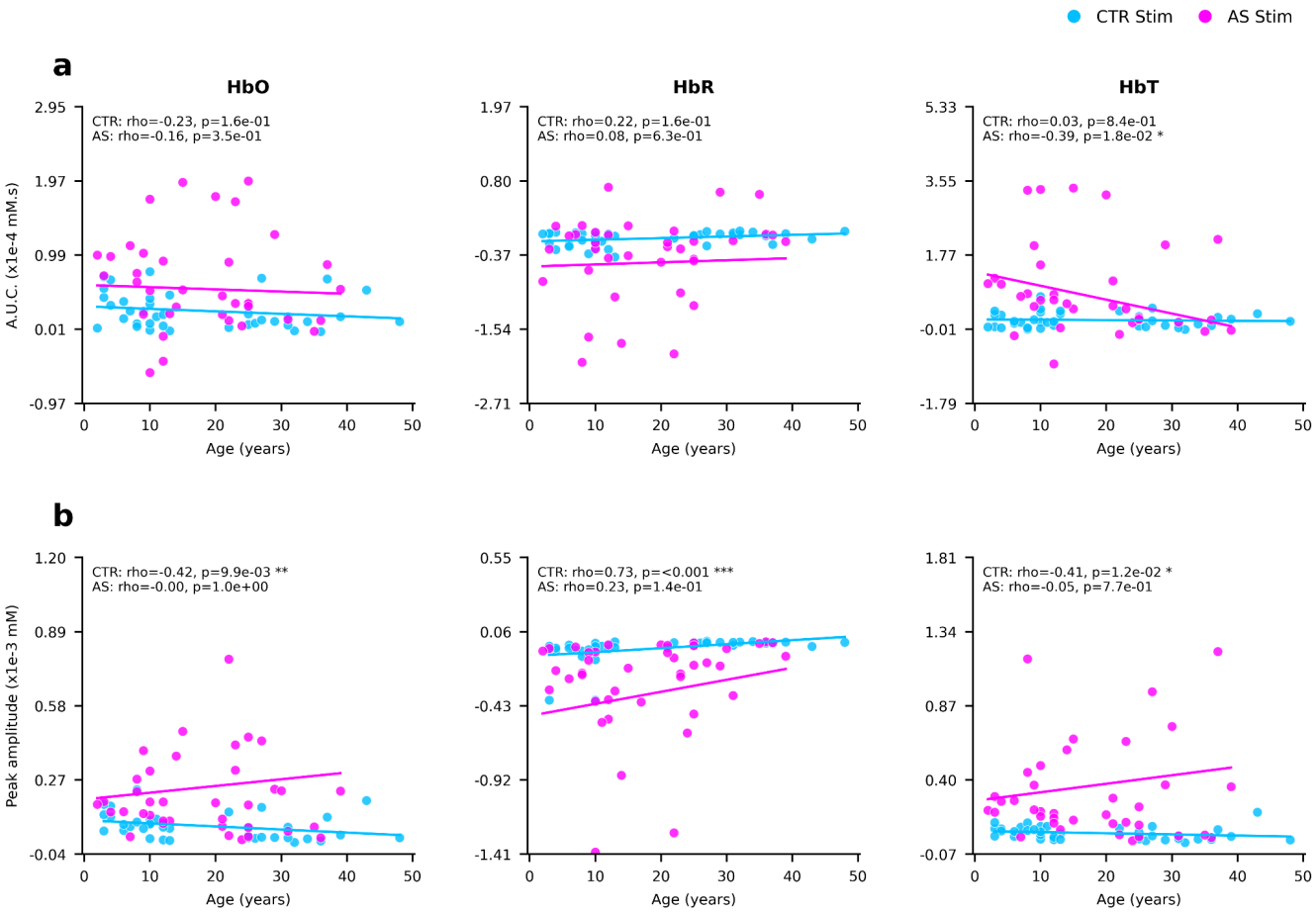
Supplementary Figure 4. Representative Western blot analysis of UBE3A expression in the cerebral cortex following therapeutic interventions. (a) Representative blot showing UBE3A protein levels in the cerebral cortex of WT non-injected mice and AS mice treated with PBS or AAV. Actin was used as a loading control. (b) Representative blot showing UBE3A expression in the cerebral cortex of WT non-injected mice and AS mice treated with PBS or ASO. α-tubulin was used as a loading control.



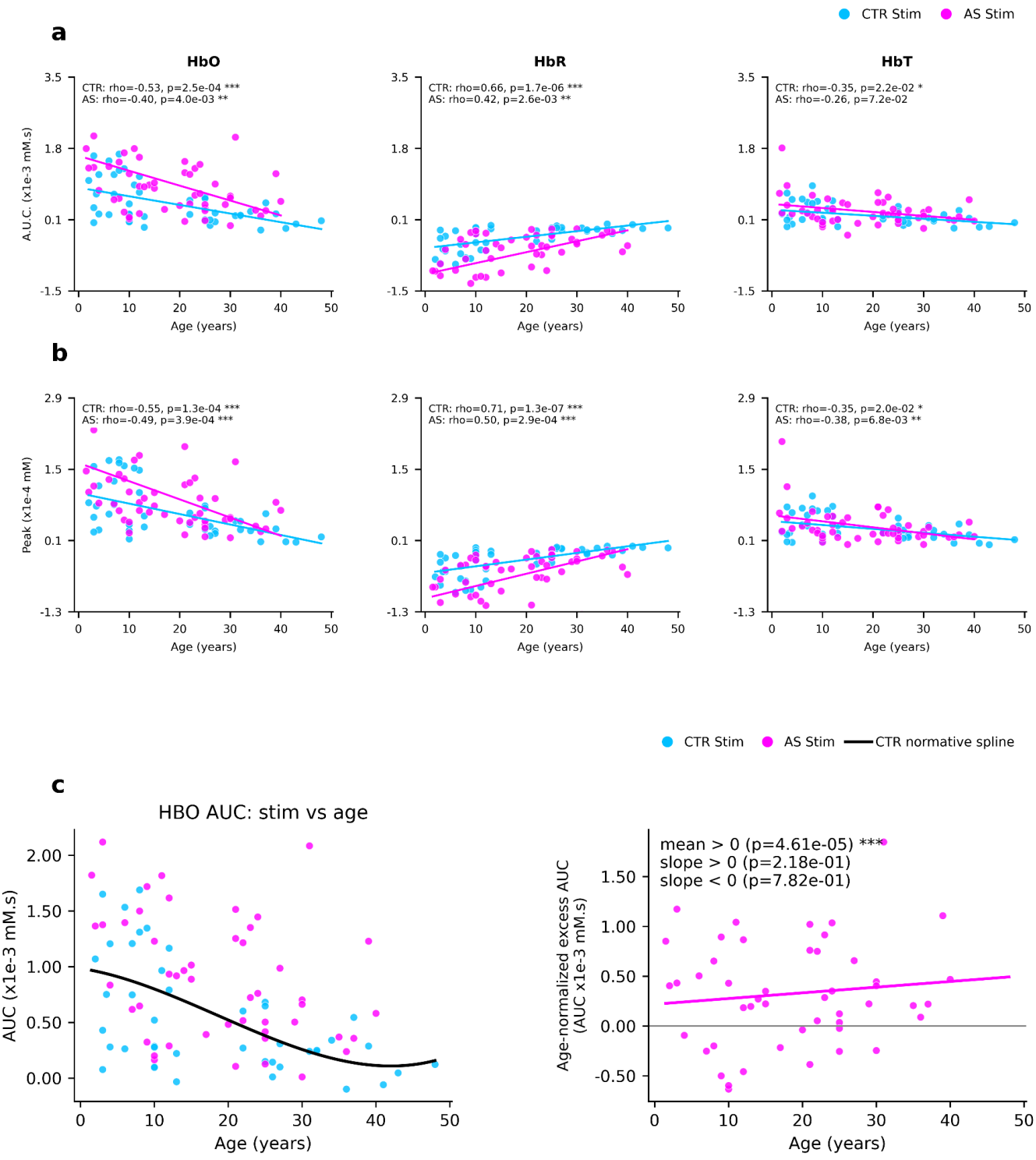
Supplementary Figure 5. Response shape conformity and temporal dynamics of the hemodynamic response to stimulation (method 2). (a) Comparison of HbO stimulation responses in AS obtained with method 1 and method 2 relative to a canonical block hemodynamic response function (HRF). The left and middle panels show normalized grand-average HbO traces for method 1 and method 2, respectively, overlaid with the canonical HRF; the right panel shows participant-level normalized root-mean-square error (RMSE), where lower values indicate closer agreement with the expected response shape. Method 2 yielded significantly lower RMSE compared to method 1 ($p < 0.001$). (b) Time to peak (TTP, seconds) during stimulation for HbO, HbR, and HbT in CTR and AS participants. Boxplots show median, interquartile range (IQR), and range; dots represent individual participants. *$p < 0.05$, **$p < 0.01$, *** $p < 0.001$.



Supplementary Figure 6. Temporal dynamics of the hemodynamic response to stimulation (method 1). Boxplots show the time to peak (seconds) for HbO, HbR, and HbT during stimulation, comparing control participants (CTR, blue) and individuals with Angelman syndrome (AS, magenta). Each point represents an individual subject; horizontal lines indicate the median, and boxes represent the interquartile range (IQR). *p < 0.05, **p < 0.01, ***p < 0.001.



Supplementary Figure 7. Age-related changes in hemodynamic response metrics (method 1). Scatterplots illustrate the relationship between participant age and hemodynamic response metrics. (a) Correlations with area under the curve (AUC), and (b) with peak amplitude. Columns correspond to hemoglobin species (HbO, HbR, HbT) and groups (CTR in blue, AS in magenta). Each point represents an individual participant, and solid lines indicate linear trends for visualization. For each panel, the sample size (n), Spearman's correlation coefficient (ρ), and p-value are reported. Statistically significant correlations are indicated by asterisks: *p < 0.05, **p < 0.01, *** p < 0.001.



Supplementary Figure 8. Age dependence and age-normalized excess response under the stricter preprocessing pipeline (method 2). (a-b) Age-related correlations for stimulation-derived area under the curve (AUC, a) and peak amplitude (b) across HbO, HbR, and HbT in CTR and AS participants. Each panel shows individual values with least-squares regression lines; statistical associations were assessed using Spearman rank correlation. (c) Age-normalized excess response analysis for HbO AUC. The left panel shows stimulation-derived HbO AUC as a function of age in CTR and AS, together with the control-derived normative

spline. The right panel shows the age-normalized excess response in AS relative to age-expected control values, with the fitted regression line. The excess response remained significantly greater than zero ($p < 0.001$), while no significant age-dependent trend was detected, indicating persistence of the pathological response without systematic increase or decline across the sampled age range. *$p < 0.05$, **$p < 0.01$, *** $p < 0.001$.

Supplementary tables

Figure	n	Stat_test	Factor	Statistic	p-value
Juvenile (P45) mice					
Visual IOS recordings					
Fig. 1c	WT n = 15 AS n = 18	Unpaired t-test		t = 3.37	0.002
Fig. 1b	WT male n = 6 WT female n = 9 AS male n = 6 AS female n = 12	Two-way ANOVA	Genotype Sex Interaction	F = 10.06 F = 0.06 F = 0.0045	0.0036 0.816 0.947
Behavioral assessment					
Supp. Fig. 1a	WT n = 15 AS n = 18	Unpaired t-test		t = 3.97	0.0004
Supp. Fig. 1b	WT n = 15 AS n = 16	Mann-Whitney test		U = 8	<0.0001
Supp. Fig. 1c	WT n = 14 AS n = 15	Unpaired t-test		t = 4.88	<0.0001
IOS – Behavior correlation in AS mice					
Supp. Fig. 2a	AS n = 18	Pearson correlation		r = - 0.15	0.542
Supp. Fig. 2b	AS n = 16	Spearman correlation		r = -0.26	0.322
Supp. Fig. 2c	AS n = 15	Pearson correlation		r = -0.003	0.992
Adult (P90) mice					
Visual IOS recordings					
Fig.1d	WT n = 15 AS n = 19	Unpaired t-test		t = 4.09	0.0003
Fig. 1e	WT male n = 7 WT female n = 8 AS male n = 11 AS female n = 8	Two-way ANOVA	Genotype Sex Interaction	F = 15.68 F = 0.16 F = 0.52	0.0004 0.688 0.477
Behavioral assessment					
Fig. 2a	WT n = 15 AS n = 17	Mann-Whitney test		U = 13	<0.0001
Fig. 2b	WT n = 15 AS n = 19	Mann-Whitney test		U = 19.5	<0.0001
Fig. 2c	WT n = 15 AS n = 15	Unpaired t-test		t = 1.84	0.077
IOS – Behavior correlation in AS mice					
Fig. 2d	AS n = 17	Spearman correlation		r = - 0.6	0.012
Fig. 2e	AS n = 19	Spearman correlation		r = -0.23	0.35
Fig. 2f	AS n = 15	Pearson correlation		r = -0.04	0.887
IOS response to treatment interventions					
Gene Therapy					
Fig. 3b	WT-NT n = 10 AS-NT n = 11 AS-PBS n =9 AS-AAV n = 11	One-Way ANOVA		F = 8.89	0.0001
Fig. 3c	WT-NT n = 9	One-Way ANOVA		F = 7.96	0.0005

	AS-NT n = 9 AS-PBS n = 8 AS-AAV n = 8			
Fig. 3d	WT-NT n = 13 AS-NT n = 13 AS-PBS n = 8 AS-AAV n = 12	One-Way ANOVA	F = 16.33	<0.0001
Fig. 3e	AS-AAV n = 8	Pearson correlation	r = -0.87	0.006
Fig. 3f	WT-NT n = 8 AS-PBS n = 4 AS-AAV n = 12	One-Way ANOVA	F = 14.53	0.0001
Fig. 3g	AS-AAV n = 8	Pearson correlation	r = -0.81	0.015
		ASO therapy		
Fig. 3i	WT-NT n = 8 AS-NT n = 11 AS-PBS n = 8 AS-ASO n = 10	One-Way ANOVA	F = 9.87	<0.0001
Fig. 3l	WT-NT n = 7 AS-PBS n = 6 AS-ASO n = 8	One-Way ANOVA	F = 43.57	<0.0001
Fig. 3m	AS-ASO n = 8	Pearson correlation	r = -0.64	0.085

Supplementary Table 1. Statistical analysis for the validation of visual IOS recordings as biomarker in the AS mouse model. Analyses were conducted in both juvenile (P45) and adult (P90) cohorts to compare IOS responses between WT and AS groups, using unpaired t-tests for pooled comparisons, and two-way ANOVA to evaluate the effects of genotype and sex. Additional analyses assessed the effects of therapeutic interventions in AS mice, including gene therapy and antisense oligonucleotide treatment, using one-way ANOVA followed by Tukey's post hoc tests for group comparisons. Associations between IOS responses, behavioral performance, and UBE3A-related measures were evaluated using correlation analyses. The table reports, for each comparison, sample size, the statistical test used, factors, test statistics, and p-values.

Category	*Parameter*	*Value*
Model information	Classifier	Random Forest
	Input features	2 IOS variables + 2 behavioural variables
	Sample size	67 mice
Overall performance	Mean accuracy	~0.88
	Accuracy (10-fold CV range)	0.84–0.90
	Full-dataset accuracy	>0.95
	AUC (ROC)	0.97
	False positive rate	0.11
	False negative rate	0.12
Cross-validation metrics (10-fold CV)	Accuracy	0.85 ± 0.12
	Precision	0.89 ± 0.17
	Recall (Sensitivity)	0.83 ± 0.17
	F1-score	0.84 ± 0.11
Feature importance	Marble (behavioural)	0.352
	Rotarod (behavioural)	0.289
	Peak Amplitude (IOS)	0.192
	Average Amplitude (IOS)	0.167

Supplementary Table 2. Random Forest classification performance and feature importance for discrimination between WT and AS mice. Summary of the Random Forest model, including overall performance, cross-validation metrics, and feature importance. The model was trained on a dataset of 67 mice and evaluated using 10-fold cross-validation. Performance is reported as accuracy range across folds, area under the receiver operating characteristic curve (AUC–ROC), and false negative rate. Cross-validation metrics are presented as mean ± standard deviation for precision and F1-score. Feature importance values indicate the relative contribution of each variable to the model, highlighting behavioral measures (rotarod performance) and intrinsic optical signal (IOS) metrics, particularly average amplitude, among the most informative predictors.

Subject ID	*Group*	*Age*	*Gender*	*Genetic Mutation*	*Movie*
AS01	AS	36	M	Deletion	Personal video
AS02	AS	21	F	Deletion	Horses
AS03	AS	22	F	Deletion	Pippi Longstocking
AS04	AS	10	M	Uniparental Disomy	Toy story
AS05	AS	3	M	Deletion	Curious George
AS06	AS	2	M	Deletion	Bing
AS07	AS	8	M	Mutation	Personal video
AS08	AS	27	F	Deletion	Harry Potter
AS09	AS	10	F	Deletion	Cinderella
AS10	AS	23	M	Deletion	Peppa Pig
AS11	AS	20	F	Mutation	Toy story 3
AS12	AS	4	F	Deletion	Masha and the Bear
AS13	AS	9	M	Deletion	Personal video
AS14	AS	7	M	Deletion	Curious George
AS15	AS	31	F	Deletion	The Hundred and One Dalmatians
AS16	AS	25	M	Uniparental Disomy	Moondance
AS17	AS	12	M	Deletion	Minions
AS18	AS	15	M	Deletion	Fantozzi clips*
AS19	AS	9	M	Deletion	Masha and the Bear
AS20	AS	6	F	Mutation	Mickey Mouse
AS21	AS	13	F	Deletion	Children playing games

AS22	AS	24	F	Deletion	Checco Zalone clips*
AS23	AS	12	M	Deletion	Personal video
AS24	AS	37	F	Mutation	Alvin 2
AS25	AS	29	F	Mutation	Alvin 2
AS26	AS	10	F	Uniparental Disomy	Bluey
AS27	AS	12	M	Deletion	Toy story
AS28	AS	39	M	Uniparental Disomy	The Christmas Chronicles 2
AS29	AS	25	F	Deletion	Harry Potter
AS30	AS	35	M	Deletion	"Ricordare" the movie**
AS31	AS	14	F	Deletion	Curious George
AS32	AS	23	M	Deletion	Pulcinella
AS33	AS	25	M	Deletion	Autobus
AS34	AS	8	M	Uniparental Disomy	"Noemi" the movie**
AS35	AS	21	F	Deletion	Personal video
AS36	AS	15	M	Deletion	Masha and the Bear
AS37	AS	22	F	Deletion	Harry Potter
AS38	AS	25	F	Deletion	Personal video
AS39	AS	30	M	Uniparental Disomy	Pope's Angelus
AS40	AS	3	M	Deletion	Bing
AS41	AS	17	F	Micro-deletion	Ben & Holly
AS42	AS	21	M	Mutation	*La stanza inclinata****

AS43	AS	11	M	Deletion	Mickey Mouse
AS44	AS	40	F	Mutation	*Stasera tutto è possibile****
AS45	AS	24	F	Deletion	Pippi Longstocking
AS46	AS	30	F	Deletion	Personal video
AS47	AS	30	M	Micro-deletion	Spirit
AS48	AS	2	F	Deletion	Masha and the Bear
AS49	AS	2	F	Deletion	Bing
CTR01	CTR	31	F	-	Futurama
CTR02	CTR	29	M	-	The Aristocats
CTR03	CTR	39	M	-	Cars
CTR04	CTR	32	F	-	Peanuts
CTR05	CTR	25	F	-	Pollon
CTR06	CTR	34	F	-	Minions
CTR07	CTR	25	M	-	The Incredibles
CTR08	CTR	26	M	-	Wall-e
CTR09	CTR	27	M	-	Finding Nemo
CTR10	CTR	31	M	-	The Simpson
CTR11	CTR	36	M	-	Peppa Pig
CTR12	CTR	4	M	-	Finding Nemo
CTR13	CTR	10	F	-	Frozen
CTR14	CTR	10	M	-	The Simpson
CTR15	CTR	3	M	-	Tom and Jerry
CTR16	CTR	13	F	-	Friends

CTR17	CTR	10	M	-	Pollon
CTR18	CTR	12	M	-	Big Hero 2
CTR19	CTR	4	F	-	Frozen
CTR20	CTR	11	M	-	The Christmas Chronicles 2
CTR21	CTR	27	F	-	Frozen
CTR22	CTR	3	F	-	Frozen
CTR23	CTR	3	M	-	Bing
CTR24	CTR	6	M	-	Gumball
CTR25	CTR	12	M	-	Harry Potter
CTR26	CTR	7	M	-	Paw Patrol
CTR27	CTR	8	M	-	Captain Tsubasa
CTR28	CTR	22	M	-	Finding Nemo
CTR29	CTR	37	F	-	*La stanza inclinata****
CTR30	CTR	43	M	-	Harry Potter
CTR31	CTR	13	F	-	Horses
CTR32	CTR	10	F	-	Lilo & Stitch
CTR33	CTR	8	M	-	Inside Out
CTR34	CTR	6	M	-	Gumball
CTR35	CTR	25	F	-	Spirit
CTR36	CTR	9	M	-	Gaming clips
CTR37	CTR	48	M	-	Peanuts
CTR38	CTR	10	M	-	Gaming clips
CTR39	CTR	2	F	-	Frozen 2

CTR40	CTR	32	F	-	Harry Potter
CTR41	CTR	7	F	-	Gumball
CTR42	CTR	3	M	-	Cars
CTR43	CTR	22	F	-	Checco Zalone clips*
CTR44	CTR	26	M	-	Checco Zalone clips*
CTR45	CTR	41	M	-	Inside Out

Supplementary Table 3. Participant demographics and experimental stimuli in AS and control groups. Summary of demographic and clinical characteristics of participants included in the study, grouped into Angelman syndrome (AS) and control (CTR) cohorts. For each subject, age (years) and sex (M, male; F, female) are reported. For AS participants, the type of genetic alteration is specified (deletion, uniparental disomy, mutation, or microdeletion). The "Movie" column indicates the visual stimulus presented during the experimental session, including standard cartoons, films, or individualized videos. *denotes a selection of films by the same director; ** indicates Italian films for which no official English translation is available; *** refers to Italian television shows. All subject identifiers are anonymized.

Metric	Unit	Hb Type	Effect	N_Subjects	N_CTR	N_AS	F	df1	df2	p_unc	η2
Method 1											
AUC	mM.s	HbO	Condition	77	39	38	18.438	1	75	0.000	0.197
AUC	mM.s	HbO	Group	77	39	38	4.877	1	75	0.030	0.061
AUC	mM.s	HbO	Group×Condition	77	39	38	0.378	1	75	0.541	0.005
AUC	mM.s	HbR	Condition	79	40	39	18.608	1	77	0.000	0.195
AUC	mM.s	HbR	Group	79	40	39	0.904	1	77	0.345	0.012
AUC	mM.s	HbR	Group×Condition	79	40	39	8.678	1	77	0.004	0.101
AUC	mM.s	HbT	Condition	75	39	36	8.557	1	73	0.005	0.105
AUC	mM.s	HbT	Group	75	39	36	9.837	1	73	0.002	0.119
AUC	mM.s	HbT	Group×Condition	75	39	36	0.996	1	73	0.321	0.013
Peak	mM	HbO	Condition	76	37	39	8.369	1	74	0.005	0.102
Peak	mM	HbO	Group	76	37	39	26.488	1	74	0.000	0.264
Peak	mM	HbO	Group×Condition	76	37	39	0.006	1	74	0.939	0.000
Peak	mM	HbR	Condition	81	38	43	18.853	1	79	0.000	0.193
Peak	mM	HbR	Group	81	38	43	37.602	1	79	0.000	0.322
Peak	mM	HbR	Group×Condition	81	38	43	13.414	1	79	0.000	0.145
Peak	mM	HbT	Condition	77	37	40	4.089	1	75	0.047	0.052
Peak	mM	HbT	Group	77	37	40	33.667	1	75	0.000	0.310
Peak	mM	HbT	Group×Condition	77	37	40	0.295	1	75	0.588	0.004
Method 2											
AUC	mM.s	HbO	Condition	94	45	49	142.0735	1	92	0.000	0.607
AUC	mM.s	HbO	Group	94	45	49	3.979	1	92	0.049	0.041
AUC	mM.s	HbO	Group×Condition	94	45	49	8.262	1	92	0.005	0.082
AUC	mM.s	HbR	Condition	94	45	49	113.235	1	92	0.000	0.552
AUC	mM.s	HbR	Group	94	45	49	3.200	1	92	0.077	0.034
AUC	mM.s	HbR	Group×Condition	94	45	49	14.767	1	92	0.000	0.138
AUC	mM.s	HbT	Condition	94	45	49	124.908	1	92	0.000	0.576
AUC	mM.s	HbT	Group	94	45	49	4.461	1	92	0.037	0.046
AUC	mM.s	HbT	Group×Condition	94	45	49	0.307	1	92	0.581	0.003
Peak	mM	HbO	Condition	94	45	49	109.548	1	92	0.000	0.544
Peak	mM	HbO	Group	94	45	49	3.014	1	92	0.086	0.032
Peak	mM	HbO	Group×Condition	94	45	49	4.948	1	92	0.029	0.051
Peak	mM	HbR	Condition	94	45	49	91.129	1	92	0.000	0.498
Peak	mM	HbR	Group	94	45	49	6.283	1	92	0.014	0.064
Peak	mM	HbR	Group×Condition	94	45	49	9.746	1	92	0.002	0.096
Peak	mM	HbT	Condition	94	45	49	114.764	1	92	0.000	0.555
Peak	mM	HbT	Group	94	45	49	0.395	1	92	0.531	0.004
Peak	mM	HbT	Group×Condition	94	45	49	0.312	1	92	0.578	0.003

Supplementary Table 4. ANOVA results for hemodynamic responses (AUC and peak) across hemoglobin species and experimental factors. Results of two-way ANOVA testing the effects of Condition, Group, and their interaction (Group × Condition) on hemodynamic response metrics for both method 1 and method 2. Analyses were performed separately for area under the curve (AUC) and peak amplitude across hemoglobin species: oxygenated (HbO), deoxygenated (HbR), and total hemoglobin (HbT). Columns report the number of subjects per group (N_Subjects, N_CTR, N_AS), F-statistics, degrees of freedom (df1, df2), uncorrected p-values (p_unc), partial eta squared (η^2) as effect size. Statistical significance was defined at $p < 0.05$ after correction. Notably, Method 1 showed a higher number of outliers, likely due to residual motion artifacts, whereas no outliers were detected using Method 2.

Metric	Unit	Hb Type	Stats_test	Comparison	N_Ctr	N_AS	T	dof	p_unc	Cohen_d	p_bh
Method 1											
AUC	mM.s	HbO	pairwise t-test	AS Mock vs AS Stim	38	38	-2.507	37.000	0.017	0.626	0.021
AUC	mM.s	HbR	pairwise t-test	AS Mock vs AS Stim	39	39	3.628	38.000	0.001	0.988	0.002
AUC	mM.s	HbT	pairwise t-test	AS Mock vs AS Stim	36	36	-1.920	35.000	0.063	0.539	0.063
AUC	mM.s	HbO	pairwise t-test	CTR Mock vs CTR Stim	39	39	-7.421	38.000	0.000	2.064	0.000
AUC	mM.s	HbR	pairwise t-test	CTR Mock vs CTR Stim	40	40	5.955	39.000	0.000	1.559	0.000
AUC	mM.s	HbT	pairwise t-test	CTR Mock vs CTR Stim	39	39	-7.470	38.000	0.000	2.122	0.000
AUC	mM.s	HbO	pairwise t-test	CTR Stim vs AS Stim	39	38	-2.303	43.913	0.026	0.531	0.029
AUC	mM.s	HbR	pairwise t-test	CTR Stim vs AS Stim	40	39	2.624	39.001	0.012	0.598	0.020
AUC	mM.s	HbT	pairwise t-test	CTR Stim vs AS Stim	39	36	-2.607	36.211	0.013	0.627	0.020
Peak	mM	HbO	pairwise t-test	AS Mock vs AS Stim	39	39	-1.473	38.000	0.149	0.261	0.168
Peak	mM	HbR	pairwise t-test	AS Mock vs AS Stim	43	43	4.145	42.000	0.000	0.707	0.000
Peak	mM	HbT	pairwise t-test	AS Mock vs AS Stim	40	40	-1.327	39.000	0.192	0.239	0.192
Peak	mM	HbO	pairwise t-test	CTR Mock vs CTR Stim	37	37	-7.689	36.000	0.000	1.578	0.000
Peak	mM	HbR	pairwise t-test	CTR Mock vs CTR Stim	38	38	3.216	37.000	0.003	0.651	0.003
Peak	mM	HbT	pairwise t-test	CTR Mock vs CTR Stim	37	37	-8.382	36.000	0.000	1.556	0.000
Peak	mM	HbO	pairwise t-test	CTR Stim vs AS Stim	37	39	-4.361	42.528	0.000	0.977	0.000
Peak	mM	HbR	pairwise t-test	CTR Stim vs AS Stim	38	43	-5.181	42.650	0.000	1.085	0.000
Peak	mM	HbT	pairwise t-test	CTR Stim vs AS Stim	37	40	-5.070	40.023	0.000	1.113	0.000
TTP	s	HbO	pairwise t-test	CTR Stim vs AS Stim	45	49	-0.957	82.983	0.342	0.194	0.512
TTP	s	HbR	pairwise t-test	CTR Stim vs AS Stim	45	49	2.601	83.561	0.011	0.543	0.033
TTP	s	HbT	pairwise t-test	CTR Stim vs AS Stim	45	49	-0.423	88.458	0.673	0.086	0.673
Method 2											
AUC	mM.s	HbO	paired t-test	AS Mock vs AS Stim	49	49	-9.627	48.000	0.000	1.920	0.000
AUC	mM.s	HbO	paired t-test	CTR Mock vs CTR Stim	45	45	-7.035	44.000	0.000	1.944	0.000
AUC	mM.s	HbR	paired t-test	AS Mock vs AS Stim	49	49	8.603	48.000	0.000	1.805	0.000
AUC	mM.s	HbR	paired t-test	CTR Mock vs CTR Stim	45	45	6.397	44.000	0.000	1.782	0.000
AUC	mM.s	HbT	paired t-test	AS Mock vs AS Stim	49	49	-8.661	48.000	0.000	1.719	0.000
AUC	mM.s	HbT	paired t-test	CTR Mock vs CTR Stim	45	45	-7.149	44.000	0.000	1.968	0.000
AUC	mM.s	HbO	Welch t-test	CTR Stim vs AS Stim	45	49	-3.059	72.032	0.003	0.616	0.004
AUC	mM.s	HbR	Welch t-test	CTR Stim vs AS Stim	45	49	3.501	63.245	0.001	0.700	0.001
AUC	mM.s	HbT	Welch t-test	CTR Stim vs AS Stim	45	49	-1.670	87.189	0.099	0.340	0.099
Peak	mM	HbO	paired t-test	AS Mock vs AS Stim	49	49	-7.805	48.000	0.000	1.226	0.000
Peak	mM	HbO	paired t-test	CTR Mock vs CTR Stim	45	45	-7.219	44.000	0.000	1.434	0.000
Peak	mM	HbR	paired t-test	AS Mock vs AS Stim	49	49	7.404	48.000	0.000	1.148	0.000
Peak	mM	HbR	paired t-test	CTR Mock vs CTR Stim	45	45	6.412	44.000	0.000	1.338	0.000
Peak	mM	HbT	paired t-test	AS Mock vs AS Stim	49	49	-7.913	48.000	0.000	1.183	0.000
Peak	mM	HbT	paired t-test	CTR Mock vs CTR Stim	45	45	-7.222	44.000	0.000	1.391	0.000
Peak	mM	HbO	Welch t-test	CTR Stim vs AS Stim	45	49	-2.074	73.405	0.042	0.418	0.047
Peak	mM	HbR	Welch t-test	CTR Stim vs AS Stim	45	49	2.932	65.442	0.005	0.587	0.006
Peak	mM	HbT	Welch t-test	CTR Stim vs AS Stim	45	49	-0.662	87.158	0.510	0.135	0.510
TTP	s	HbO	Welch t-test	CTR Stim vs AS Stim	45	49	-2.456	90.874	0.016	0.503	0.024
TTP	s	HbR	Welch t-test	CTR Stim vs AS Stim	45	49	-2.879	91.822	0.005	0.591	0.015
TTP	s	HbT	Welch t-test	CTR Stim vs AS Stim	45	49	-1.882	88.077	0.063	0.384	0.063

Supplementary Table 5. Paired t-test results comparing hemodynamic responses between experimental conditions (methods 1 and 2). Results of paired t-tests comparing hemodynamic responses between mock and stimulation conditions. Analyses were performed separately for area under the curve (AUC) and peak amplitude across hemoglobin species: oxygenated (HbO), deoxygenated (HbR), and total hemoglobin (HbT). The table reports the number of subjects (N), t-statistics (t), degrees of freedom (dof), uncorrected p-values (p_unc), effect sizes (e.g., Cohen's d), and Benjamini–Hochberg corrected p-values (p_bh) for multiple comparisons. Statistical significance was defined at $p < 0.05$ after correction. Notably, a higher number of outliers was observed in method 1, likely reflecting residual motion artifacts, whereas no outliers were detected using method 2.

References

1. Hernandez, S. M. & Pollonini, L. NIRSplot: A Tool for Quality Assessment of fNIRS Scans. in *Biophotonics Congress: Biomedical Optics 2020 (Translational, Microscopy, OCT, OTS, BRAIN)* BM2C.5 (Optica Publishing Group, Washington, DC, 2020). doi:10.1364/BRAIN.2020.BM2C.5.
2. Gramfort, A. *et al.* MEG and EEG data analysis with MNE-Python. *Front Neurosci* **7**, 267 (2013).
3. Luke, R. *et al.* Analysis methods for measuring passive auditory fNIRS responses generated by a block-design paradigm. *Neurophotonics* **8**, 025008 (2021).
4. Huppert, T. J., Diamond, S. G., Franceschini, M. A. & Boas, D. A. HomER: a review of time-series analysis methods for near-infrared spectroscopy of the brain. *Appl Opt* **48**, D280-298 (2009).
5. Cui, X., Bray, S. & Reiss, A. L. Functional near infrared spectroscopy (NIRS) signal improvement based on negative correlation between oxygenated and deoxygenated hemoglobin dynamics. *Neuroimage* **49**, 3039–3046 (2010).
6. Liu, Y., Sánchez Hernández, F., Ting, F. & Hyde, D. C. Comparing fixed-array and functionally-defined channel of interest approaches to infant functional near-infrared spectroscopy data. *Neuroimage* **261**, 119520 (2022).