

# Supplementary materials

It takes two: digitally assisted differential diagnostics of ASD based on dyadic interactions

I S Plank et al.

2025-07-24

## Contents

<b>1</b>	<b>S1 Version numbers for R and all packages</b>	<b>1</b>
<b>2</b>	<b>S2 Inclusion and exclusion criteria</b>	<b>2</b>
<b>3</b>	<b>S3 Measures</b>	<b>2</b>
3.1	S3.1 Group comparisons based on diagnostic status . . . . .	2
<b>4</b>	<b>S4 Features</b>	<b>3</b>
4.1	S4.1 Facial expressions extracted from OpenFace . . . . .	3
4.2	S4.2 Motion quantity extracted using Motion Energy Analysis . . . . .	4
4.3	S4.3 Speech and turn-taking features . . . . .	4
4.4	S4.4 Cross-modal features . . . . .	5
4.5	S4.5 Synchrony computations . . . . .	5
4.6	S4.6 Feature lists . . . . .	5
<b>5</b>	<b>S5 Model performance</b>	<b>8</b>
5.1	S5.1 Distinguishing BPD-involved from COMP interactions . . . . .	8
5.2	S5.2 One-versus-One comparisons . . . . .	8
5.3	S5.3 Multi-group comparisons . . . . .	9
5.4	S5.4 Gender comparisons . . . . .	9
<b>6</b>	<b>S6 Model visualisations</b>	<b>10</b>
6.1	S6.1 ASD-involved versus BPD-involved classifiers . . . . .	10
6.2	S6.2 ASD-involved versus COMP classifiers . . . . .	14
6.3	S6.3 ASD-involved versus COMP classifiers . . . . .	18
<b>7</b>	<b>S7 References</b>	<b>21</b>

## 1 S1 Version numbers for R and all packages

```
## [1] "R version 4.5.1 (2025-06-13)"
## [1] "tidyverse version 2.0.0"
## [1] "knitr version 1.50"
## [1] "ggpubr version 0.6.1"
## [1] "gggrain version 0.0.4"
## [1] "Hmisc version 5.2.3"
## [1] "rstatix version 0.7.2"
## [1] "emmeans version 1.11.2"
## [1] "flextable version 0.9.9"
## [1] "officer version 0.6.10"
## [1] "english version 1.2.6"
```

## 2 S2 Inclusion and exclusion criteria

Inclusion criteria applying to all participants:

- no neurological diagnoses
- at least 18 and at most 60 years of age
- normal or corrected-to-normal vision
- legally binding, written informed consent from the participant for study participation

Additional inclusion criteria for ASD participants:

- valid F84 diagnosis according to ICD-10
- normal speaking abilities

Additional inclusion criteria for BPD participants:

- valid F60.31 diagnosis according to ICD-10
- no F84 diagnosis according to ICD-10
- IQ estimate above 70

Additional inclusion criteria for comparison participants:

- IQ estimate above 70
- no psychiatric diagnoses
- no current psychotropic medication

Furthermore, we applied the following exclusion criteria:

- underage people or people older than 60 years
- inability to speak
- current or previous neurological disorder
- acute suicidality, self-harming or aggressive behaviour
- lack of legally binding informed consent

Testing was discontinued if participants withdrew their consent, exclusion criteria applied or in the case of technical difficulties.

## 3 S3 Measures

We collected the following self-report questionnaires:

- Autism-like traits: Autism Quotient, AQ<sup>1</sup>
- Empathy: Saarbrücker Persönlichkeitsfragebogen, SPF<sup>2</sup>, which is the German version of the interpersonal reactivity index
- Alexithymia: Toronto Alexithymia Scale, TAS20<sup>3</sup>
- Depressiveness: Beck Depression Inventory, BDI<sup>4</sup>
- Self-monitoring: Self-monitoring Scale, SMS<sup>5</sup>
- Movement difficulties: Adult Dyspraxia Checklist, ADC<sup>6</sup>

For all participants recruited for this project, we also collected the short version of the Borderline Symptom List, BSL-23<sup>7</sup>.

For 37 of the dyads, we used a plexiglass screen placed between the interaction partners on the table to decrease the chances of spreading an undetected infection. We applied a transparent anti-reflection foil to reduce any mirroring effects. During the conversation, participants took off their face masks. We asked these participants how much the plexiglass influenced them during the interactions (plexi; scale from 0 to 3). All participants were asked how much the cameras influenced their behaviour (video; scale from 0 to 3). Rapport is a sum of the following ratings, all on scales from 0 to 3, thus ranging from 0 to 15:

- How likeable was your interaction partner?
- How friendly was your interaction partner?
- How comfortable did you feel during the conversation?
- How smooth was the communication between you and your conversation partner?
- How well did your conversation partner respond to you?

### 3.1 S3.1 Group comparisons based on diagnostic status

measurement	ASD	BPD	COMP	BPDvsASD	COMPvsASD	COMPvsBPD
ADC_total	49.06 ( $\pm 17.17$ ), n = 17	76.95 ( $\pm 14.98$ ), n = 21	34.99 ( $\pm 23.86$ ), n = 82	0.006*	0.243	0.000*
AQ_total	34.29 ( $\pm 6.18$ ), n = 17	23.67 ( $\pm 4.72$ ), n = 21	14.57 ( $\pm 5.27$ ), n = 82	0.000*	0.000*	0.000*
BDI_total	15.94 ( $\pm 11.57$ ), n = 17	24.81 ( $\pm 10.48$ ), n = 21	3.99 ( $\pm 3.79$ ), n = 82	0.715	0.000*	0.000*
BERT_acc	0.80 ( $\pm 0.07$ ), n = 17	0.82 ( $\pm 0.09$ ), n = 21	0.83 ( $\pm 0.07$ ), n = 82	1.000	0.919	1.000
BERT_rt	5.79 ( $\pm 2.78$ ), n = 17	3.84 ( $\pm 1.24$ ), n = 21	3.33 ( $\pm 1.33$ ), n = 82	0.453	0.000*	0.637
BSL_total	NaN ( $\pm$ NA), n = 0	46.71 ( $\pm 21.44$ ), n = 21	7.19 ( $\pm 6.84$ ), n = 37	NA	NA	0.000*
IQ_estimate	116.47 ( $\pm 13.90$ ), n = 17	109.60 ( $\pm 9.54$ ), n = 20	112.77 ( $\pm 13.26$ ), n = 82	1.000	1.000	1.000
SMS_total	5.47 ( $\pm 2.81$ ), n = 17	11.67 ( $\pm 2.11$ ), n = 21	9.80 ( $\pm 2.62$ ), n = 82	0.000*	0.000*	0.027*
SPF_total	37.00 ( $\pm 7.20$ ), n = 17	40.95 ( $\pm 8.36$ ), n = 21	44.73 ( $\pm 5.82$ ), n = 82	1.000	0.000*	0.328
TAS_total	61.76 ( $\pm 11.03$ ), n = 17	55.33 ( $\pm 11.60$ ), n = 21	37.55 ( $\pm 8.53$ ), n = 82	1.000	0.000*	0.000*
age	37.59 ( $\pm 13.19$ ), n = 17	28.38 ( $\pm 10.02$ ), n = 21	28.89 ( $\pm 10.18$ ), n = 82	0.506	0.243	1.000
plexi	0.82 ( $\pm 0.53$ ), n = 17	0.85 ( $\pm 0.90$ ), n = 13	0.86 ( $\pm 0.60$ ), n = 43	1.000	1.000	1.000
rapport	12.24 ( $\pm 2.19$ ), n = 17	11.95 ( $\pm 2.64$ ), n = 21	12.33 ( $\pm 2.39$ ), n = 81	1.000	1.000	1.000
video	0.71 ( $\pm 0.69$ ), n = 17	0.81 ( $\pm 0.75$ ), n = 21	0.63 ( $\pm 0.68$ ), n = 82	1.000	1.000	1.000

```
##          ilabel
## gender ASD BPD COMP
##      fem   6  14  52
##      mal  11   7  30

##
## Pearson's Chi-squared test
##
## data:  tb.gen
## X-squared = 5.1108, df = 2, p-value = 0.07766
```

## 4 S4 Features

### 4.1 S4.1 Facial expressions extracted from OpenFace

We only included data of participants with a mean confidence of tracked frames greater than 75% and more than 90% successfully tracked frames. Facial expressions were captured as action units. We did not extract emotional expressions from these facial expressions as coherence between facial expressions and emotions is not a given and might be even less so for autistic people<sup>8</sup>.

For the calculation of synchronisation, we included rotational parameters (yaw, roll, pitch) as well as the same action units as in our previous study<sup>9</sup>:

- Mealplanning: 1, 2, 6, 7, 9, 14, 15, 17, 20, 25, 26 and 45
- Hobbies: 1, 2, 6, 7, 9, 15, 17, 20, 23, 25, 26 and 45

We also extracted total facial expressiveness as mean intensity of all action units for each interaction partner to be included in the *MovEx* and the *CROSSturn* models. For the *CROSSturn* model, we also included other action units, as listed below.

These correspond to the following movements:

- AU1: inner brow raiser
- AU2: outer brow raiser
- AU4: brow lowerer (only *CROSSturn*)
- AU5: upper lid raiser (only *CROSSturn*)
- AU6: cheek raiser
- AU7: lid tightener
- AU9: nose wrinkler
- AU10: upper lip raiser (only *CROSSturn*)
- AU12: lip corner puller (only *CROSSturn*)

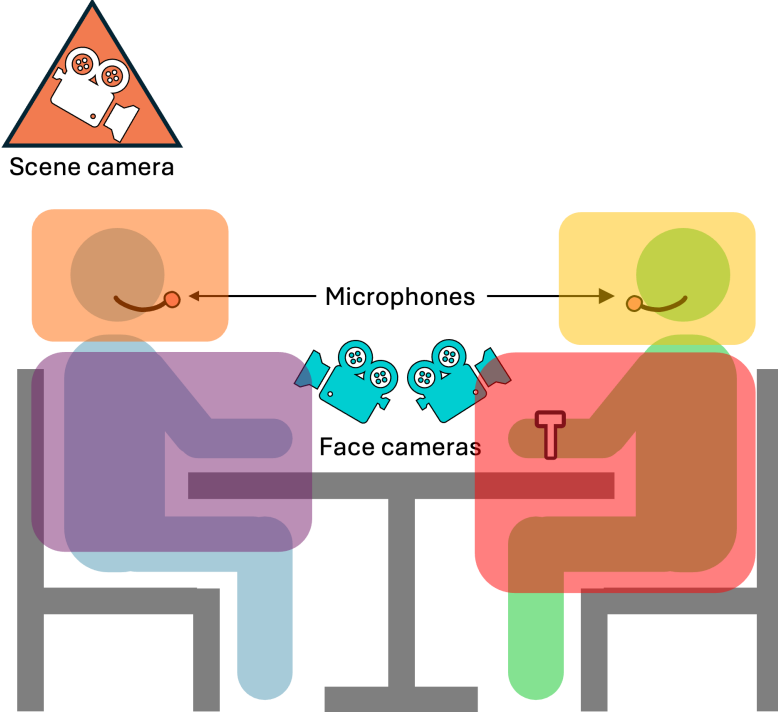
- AU14: dimpler
- AU15: lip corner depressor
- AU17: chin raiser
- AU20: lip stretcher
- AU23: lip tightener
- AU25: lips part
- AU26: jaw drop
- AU45: blink

Furthermore, we used translational head position parameters to infer head motion using the following formula with  $\Delta_t$  referring to the respective frame-to-frame changes:

$$\text{head movement} = \sqrt{\Delta_t x^2 + \Delta_t y^2 + \Delta_t z^2}$$

## 4.2 S4.2 Motion quantity extracted using Motion Energy Analysis

This figure shows body (red and purple) and head (yellow and orange) regions of interests of each interaction partner separately:



There was always a space between the head and body region. Regions were chosen such that they cover the full range of motion throughout one conversation of one interaction partner. Thus, their sizes differed which is why we scaled all values.

In addition to using the motion quantity to compute synchronisation, we also extracted total movement in each region of interest for each interaction partner to be included in the *MovEx* model.

## 4.3 S4.3 Speech and turn-taking features

We extracted pitch using praat’s autocorrelation method, a technique widely recognized for its reliability and accuracy<sup>10</sup>. We implemented a two-step pitch extraction method, as outlined by Hirst<sup>11</sup>. First, to capture a broad range of frequencies, we set a low pitch floor of 50 Hz and a high pitch ceiling of 700 Hz, with a time step of 15 ms. All other parameters were set to Praat’s default values. Second, using these initial pitch values, we determined the first and third quartiles of pitch for each participant and task. We then used these quartiles to compute individual pitch floors and ceilings with the following algorithm:

$$\text{floor} = \min(0.75 \cdot Q_{1,hobbies}, 0.75 \cdot Q_{1,mealplanning})$$

$$\text{ceiling} = \max(2.5 \cdot Q_{3,hobbies}, 2.5 \cdot Q_{3,mealplanning})$$

We then used these individual pitch floors (range = 46 to 168Hz, mean =  $109.7 \pm 31.8$ ) and ceilings (range = 250 to 839Hz, mean =  $481.7 \pm 147.4$ ) to extract pitch. To ensure an equal number of frames for all participants, we maintained a consistent time step across all analyses. By default, praat calculates this time step using the following formula:

$$\text{timestep} = \frac{0.75}{\text{floor}}$$

Here, we used the same time step of 0.016 as in our previous study<sup>12</sup> which was determined based on the minimum individual pitch floor of that sample. Since the new sample did not include anyone with a lower pitch floor, the time step fits both samples.

Intensity was extracted by convolving the squared sound with a Gaussian analysis window. We used praat’s default values of minimum pitch 100Hz and time step of 0.01s.

To estimate synchrony, we extracted continuous pitch and intensity time series for every millisecond of the recording. For pitch extraction, we used consistent parameters across all participants instead of individualized settings. This was necessary because the analysis width depends on the pitch floor. Given the heterogeneity of our sample, we opted for a wide range of considered frequencies, setting the pitch floor at 50Hz and the pitch ceiling at 700Hz. For intensity, we relied on Praat’s default values.

In the case of turn-based synchronization, we correlated the median pitch or intensity of each turn with the median pitch or intensity of the preceding turn.

Next, we used the uhm-o-meter<sup>13,14</sup> to differentiate between periods of speaking and silence, identify syllables and extract several prosodic features (total number of syllables, total number of silent phases, duration of speaking as phonation time, speech rate as number of syllables per second, articulation rate as number of syllables per phonation time, average syllable duration and silence-to-turn ratio). The resulting speaking and silent instances were visually and aurally inspected to verify the accuracy of the algorithm.

#### 4.4 S4.4 Cross-modal features

We captured two types of cross-modal features:

- Interpersonal synchronisation of one person’s head with the other person’s body movement and vice versa
- AU activation, body and head movement during listening and speaking

#### 4.5 S4.5 Synchrony computations

We used the following settings for our windowed lagged cross-correlation (WLCC):

Table 2: WLCC settings in seconds

measure	window	step	lag
Facial action units synchronisation	7	4	2
Body MEA synchronisation	30	15	5
Head MEA synchronisation	30	15	5
Intrapersonal synchrony	30	15	5
Pitch synchrony	16	8	2
Intensity synchrony	16	8	2

For each window, the maximum correlation value was chosen out of all relevant lags (peak-picking). We cross-correlated head movements (from OpenFace) with body motion energy time series (from MEA) to estimate intrapersonal synchrony.

#### 4.6 S4.6 Feature lists

This list shows all features without the information of conversation, i.e., each of these features was added twice to the model, once from the mealplanning and once from the hobbies conversation. The number of features per model is displayed as well. For many of the extracted features, we calculated summary scores some which are indicated by abbreviations (mean, md = median, sd = standard deviation, min = minimum, max = maximum, ske = skewness and kurtosis)

features	no of fea- tures
BODYSYNC: min_M_bodysync, max_M_bodysync, sd_M_bodysync, mean_M_bodysync, md_M_bodysync, skew_M_bodysync, kurtosis_M_bodysync	14
CROSSLOC: min_M_LOF, min_M_ROF, max_M_LOF, max_M_ROF, md_M_LOF, md_M_ROF, mean_M_LOF, mean_M_ROF, sd_M_LOF, sd_M_ROF, kurtosis_M_LOF, kurtosis_M_ROF, skew_M_LOF, skew_M_ROF	28

features	no of fea- tures
CRF_min_M_AU01_r, self_min_M_AU02_r, self_min_M_AU04_r, self_min_M_AU05_r, self_min_M_AU06_r, self_min_M_AU07_r, self_min_M_AU09_r, self_min_M_AU10_r, self_min_M_AU12_r, self_min_M_AU14_r, self_min_M_AU15_r, self_min_M_AU17_r, self_min_M_AU20_r, self_min_M_AU23_r, self_min_M_AU25_r, self_min_M_AU26_r, self_min_M_AU45_r, self_min_M_MEA_body, self_min_M_MEA_head, self_max_M_AU01_r, self_max_M_AU02_r, self_max_M_AU04_r, self_max_M_AU05_r, self_max_M_AU06_r, self_max_M_AU07_r, self_max_M_AU09_r, self_max_M_AU10_r, self_max_M_AU12_r, self_max_M_AU14_r, self_max_M_AU15_r, self_max_M_AU17_r, self_max_M_AU20_r, self_max_M_AU23_r, self_max_M_AU25_r, self_max_M_AU26_r, self_max_M_AU45_r, self_max_M_MEA_body, self_max_M_MEA_head, self_md_M_AU01_r, self_md_M_AU02_r, self_md_M_AU04_r, self_md_M_AU05_r, self_md_M_AU06_r, self_md_M_AU07_r, self_md_M_AU09_r, self_md_M_AU10_r, self_md_M_AU12_r, self_md_M_AU14_r, self_md_M_AU15_r, self_md_M_AU17_r, self_md_M_AU20_r, self_md_M_AU23_r, self_md_M_AU25_r, self_md_M_AU26_r, self_md_M_AU45_r, self_md_M_MEA_body, self_md_M_MEA_head, self_mean_M_AU01_r, self_mean_M_AU02_r, self_mean_M_AU04_r, self_mean_M_AU05_r, self_mean_M_AU06_r, self_mean_M_AU07_r, self_mean_M_AU09_r, self_mean_M_AU10_r, self_mean_M_AU12_r, self_mean_M_AU14_r, self_mean_M_AU15_r, self_mean_M_AU17_r, self_mean_M_AU20_r, self_mean_M_AU23_r, self_mean_M_AU25_r, self_mean_M_AU26_r, self_mean_M_AU45_r, self_mean_M_MEA_body, self_mean_M_MEA_head, self_sd_M_AU01_r, self_sd_M_AU02_r, self_sd_M_AU04_r, self_sd_M_AU05_r, self_sd_M_AU06_r, self_sd_M_AU07_r, self_sd_M_AU09_r, self_sd_M_AU10_r, self_sd_M_AU12_r, self_sd_M_AU14_r, self_sd_M_AU15_r, self_sd_M_AU17_r, self_sd_M_AU20_r, self_sd_M_AU23_r, self_sd_M_AU25_r, self_sd_M_AU26_r, self_sd_M_AU45_r, self_sd_M_MEA_body, self_sd_M_MEA_head, self_kurtosis_M_AU01_r, self_kurtosis_M_AU02_r, self_kurtosis_M_AU04_r, self_kurtosis_M_AU05_r, self_kurtosis_M_AU06_r, self_kurtosis_M_AU07_r, self_kurtosis_M_AU09_r, self_kurtosis_M_AU10_r, self_kurtosis_M_AU12_r, self_kurtosis_M_AU14_r, self_kurtosis_M_AU15_r, self_kurtosis_M_AU17_r, self_kurtosis_M_AU20_r, self_kurtosis_M_AU23_r, self_kurtosis_M_AU25_r, self_kurtosis_M_AU26_r, self_kurtosis_M_AU45_r, self_kurtosis_M_MEA_body, self_kurtosis_M_MEA_head, self_skew_M_AU01_r, self_skew_M_AU02_r, self_skew_M_AU04_r, self_skew_M_AU05_r, self_skew_M_AU06_r, self_skew_M_AU07_r, self_skew_M_AU09_r, self_skew_M_AU10_r, self_skew_M_AU12_r, self_skew_M_AU14_r, self_skew_M_AU15_r, self_skew_M_AU17_r, self_skew_M_AU20_r, self_skew_M_AU23_r, self_skew_M_AU25_r, self_skew_M_AU26_r, self_skew_M_AU45_r, self_skew_M_MEA_body, self_skew_M_MEA_head, other_min_M_AU01_r, other_min_M_AU02_r, other_min_M_AU04_r, other_min_M_AU05_r, other_min_M_AU06_r, other_min_M_AU07_r, other_min_M_AU09_r, other_min_M_AU10_r, other_min_M_AU12_r, other_min_M_AU14_r, other_min_M_AU15_r, other_min_M_AU17_r, other_min_M_AU20_r, other_min_M_AU23_r, other_min_M_AU25_r, other_min_M_AU26_r, other_min_M_AU45_r, other_min_M_MEA_body, other_min_M_MEA_head, other_max_M_AU01_r, other_max_M_AU02_r, other_max_M_AU04_r, other_max_M_AU05_r, other_max_M_AU06_r, other_max_M_AU07_r, other_max_M_AU09_r, other_max_M_AU10_r, other_max_M_AU12_r, other_max_M_AU14_r, other_max_M_AU15_r, other_max_M_AU17_r, other_max_M_AU20_r, other_max_M_AU23_r, other_max_M_AU25_r, other_max_M_AU26_r, other_max_M_AU45_r, other_max_M_MEA_body, other_max_M_MEA_head, other_md_M_AU01_r, other_md_M_AU02_r, other_md_M_AU04_r, other_md_M_AU05_r, other_md_M_AU06_r, other_md_M_AU07_r, other_md_M_AU09_r, other_md_M_AU10_r, other_md_M_AU12_r, other_md_M_AU14_r, other_md_M_AU15_r, other_md_M_AU17_r, other_md_M_AU20_r, other_md_M_AU23_r, other_md_M_AU25_r, other_md_M_AU26_r, other_md_M_AU45_r, other_md_M_MEA_body, other_md_M_MEA_head, other_mean_M_AU01_r, other_mean_M_AU02_r, other_mean_M_AU04_r, other_mean_M_AU05_r, other_mean_M_AU06_r, other_mean_M_AU07_r, other_mean_M_AU09_r, other_mean_M_AU10_r, other_mean_M_AU12_r, other_mean_M_AU14_r, other_mean_M_AU15_r, other_mean_M_AU17_r, other_mean_M_AU20_r, other_mean_M_AU23_r, other_mean_M_AU25_r, other_mean_M_AU26_r, other_mean_M_AU45_r, other_mean_M_MEA_body, other_mean_M_MEA_head, other_sd_M_AU01_r, other_sd_M_AU02_r, other_sd_M_AU04_r, other_sd_M_AU05_r, other_sd_M_AU06_r, other_sd_M_AU07_r, other_sd_M_AU09_r, other_sd_M_AU10_r, other_sd_M_AU12_r, other_sd_M_AU14_r, other_sd_M_AU15_r, other_sd_M_AU17_r, other_sd_M_AU20_r, other_sd_M_AU23_r, other_sd_M_AU25_r, other_sd_M_AU26_r, other_sd_M_AU45_r, other_sd_M_MEA_body, other_sd_M_MEA_head, other_kurtosis_M_AU01_r, other_kurtosis_M_AU02_r, other_kurtosis_M_AU04_r, other_kurtosis_M_AU05_r, other_kurtosis_M_AU06_r, other_kurtosis_M_AU07_r, other_kurtosis_M_AU09_r, other_kurtosis_M_AU10_r, other_kurtosis_M_AU12_r, other_kurtosis_M_AU14_r, other_kurtosis_M_AU15_r, other_kurtosis_M_AU17_r, other_kurtosis_M_AU20_r, other_kurtosis_M_AU23_r, other_kurtosis_M_AU25_r, other_kurtosis_M_AU26_r, other_kurtosis_M_AU45_r, other_kurtosis_M_MEA_body, other_kurtosis_M_MEA_head,	532

features	no of fea- tures
<i>CROSSsync</i> , max_M_AU01_r, sd_M_AU01_r, mean_M_AU01_r, md_M_AU01_r, skew_M_AU01_r, kurtosis_M_AU01_r, min_M_AU02_r, max_M_AU02_r, sd_M_AU02_r, mean_M_AU02_r, md_M_AU02_r, skew_M_AU02_r, kurtosis_M_AU02_r, min_M_AU06_r, max_M_AU06_r, sd_M_AU06_r, mean_M_AU06_r, md_M_AU06_r, skew_M_AU06_r, kurtosis_M_AU06_r, min_M_AU07_r, max_M_AU07_r, sd_M_AU07_r, mean_M_AU07_r, md_M_AU07_r, skew_M_AU07_r, kurtosis_M_AU07_r, min_M_AU09_r, max_M_AU09_r, sd_M_AU09_r, mean_M_AU09_r, md_M_AU09_r, skew_M_AU09_r, kurtosis_M_AU09_r, min_M_AU14_r, max_M_AU14_r, sd_M_AU14_r, mean_M_AU14_r, md_M_AU14_r, skew_M_AU14_r, kurtosis_M_AU14_r, min_M_AU15_r, max_M_AU15_r, sd_M_AU15_r, mean_M_AU15_r, md_M_AU15_r, skew_M_AU15_r, kurtosis_M_AU15_r, min_M_AU17_r, max_M_AU17_r, sd_M_AU17_r, mean_M_AU17_r, md_M_AU17_r, skew_M_AU17_r, kurtosis_M_AU17_r, min_M_AU20_r, max_M_AU20_r, sd_M_AU20_r, mean_M_AU20_r, md_M_AU20_r, skew_M_AU20_r, kurtosis_M_AU20_r, min_M_AU25_r, max_M_AU25_r, sd_M_AU25_r, mean_M_AU25_r, md_M_AU25_r, skew_M_AU25_r, kurtosis_M_AU25_r, min_M_AU26_r, max_M_AU26_r, sd_M_AU26_r, mean_M_AU26_r, md_M_AU26_r, skew_M_AU26_r, kurtosis_M_AU26_r, min_M_AU45_r, max_M_AU45_r, sd_M_AU45_r, mean_M_AU45_r, md_M_AU45_r, skew_M_AU45_r, kurtosis_M_AU45_r	168
<i>HEADsync</i> , max_M_headsync, sd_M_headsync, mean_M_headsync, md_M_headsync, skew_M_headsync, kurtosis_M_headsync, min_M_pose_Rxsync, max_M_pose_Rxsync, sd_M_pose_Rxsync, mean_M_pose_Rxsync, md_M_pose_Rxsync, skew_M_pose_Rxsync, kurtosis_M_pose_Rxsync, min_M_pose_Rysync, max_M_pose_Rysync, sd_M_pose_Rysync, mean_M_pose_Rysync, md_M_pose_Rysync, skew_M_pose_Rysync, kurtosis_M_pose_Rysync, min_M_pose_Rzsync, max_M_pose_Rzsync, sd_M_pose_Rzsync, mean_M_pose_Rzsync, md_M_pose_Rzsync, skew_M_pose_Rzsync, kurtosis_M_pose_Rzsync	56
<i>INTRAsync</i> , max_M_intra, sd_M_intra, mean_M_intra, md_M_intra, skew_M_intra, kurtosis_M_intra	14
<i>MovEx</i> , total_movement, M_head_total_movement, mean_intensity_M	6
<i>Speech</i> , dyad_int_sync_MEA_M_speech, dyad_spr_M_speech, dyad_str_M_speech, dyad_ttg_M_speech, dyad_no_turns_M_speech, nsyll_M_speech, npause_M_speech, pho_M_speech, art_M_speech, pit_sync_M_speech, int_sync_M_speech, art_sync_M_speech, pit_var_M_speech, int_var_M_speech	30

## 5 S5 Model performance

### 5.1 S5.1 Distinguishing BPD-involved from COMP interactions

While developing an algorithm for technology-assisted diagnostics of BPD was not the explicit goal of this research project, we explored the application of our features to the classification between BPD-involved and COMP interactions. Despite the features being chosen with symptoms and characteristics of ASD in mind, the *CROSSsturn*, *FACEsync*, *HEADsync* and *Speech* models performed above chance in this comparison (*BODYsync*:  $p_{FDR} = 1$ ; *CROSSsync*:  $p_{FDR} = 0.282$ ; *INTRAsync*:  $p_{FDR} = 1$ ; *MovEx*:  $p_{FDR} = 0.242$ ). Specifically, the *HEADsync* model achieved 68.7% balanced accuracy (71.4; 65.9% specificity), the *FACEsync* 64% (64.3; 65.9% specificity), the *Speech* 61.6% (59.5; 63.6% specificity) and the *CROSSsturn* 55.7% (52.4; 59.1% specificity). The stacking model performed comparable to the *HEADsync* and the *MovEx* model but outperformed the other base models (see [!T]), reaching 65.3% balanced accuracy (73.8; 56.8% specificity). Thus, the stacking model only misclassified eleven BPD-involved interactions as COMP, but 19 COMP interactions were labelled as BPD-involved.

### 5.2 S5.2 One-verus-One comparisons

comparison	model	sens	spec	BAC	AUC	p.fdr	sig
ASD-COMP vs BPD-COMP	<i>BODYsync</i>	32.353	47.619	39.986	0.361	1.000	
ASD-COMP vs BPD-COMP	<i>CROSSsync</i>	52.941	69.048	60.994	0.678	0.136	
ASD-COMP vs BPD-COMP	<i>CROSSsturn</i>	67.647	85.714	76.681	0.749	0.000	*
ASD-COMP vs BPD-COMP	<i>FACEsync</i>	76.471	59.524	67.997	0.719	0.000	*
ASD-COMP vs BPD-COMP	<i>HEADsync</i>	55.882	57.143	56.513	0.609	0.936	
ASD-COMP vs BPD-COMP	<i>INTRAsync</i>	38.235	54.762	46.499	0.455	1.000	
ASD-COMP vs BPD-COMP	<i>MovEx</i>	58.824	78.571	68.698	0.758	0.027	*
ASD-COMP vs BPD-COMP	<i>Speech</i>	64.706	76.190	70.448	0.771	0.000	*
ASD-COMP vs BPD-COMP	STACK	70.588	92.857	81.723	0.805	NaN	NA
ASD-COMP vs COMP-COMP	<i>BODYsync</i>	58.824	61.364	60.094	0.607	0.000	*



comparison	model	sens	spec	BAC	AUC	p.fdr	sig
ASD-COMP vs COMP-COMP	CROSSsync	67.647	75.000	71.324	0.763	0.000	*
ASD-COMP vs COMP-COMP	CROSSturn	44.118	72.727	58.422	0.627	0.000	*
ASD-COMP vs COMP-COMP	FACEsync	79.412	70.454	74.933	0.780	0.000	*
ASD-COMP vs COMP-COMP	HEADsync	58.824	68.182	63.503	0.608	1.000	
ASD-COMP vs COMP-COMP	INTRAsync	26.471	40.909	33.690	0.297	1.000	
ASD-COMP vs COMP-COMP	MovEx	64.706	72.727	68.717	0.777	0.000	*
ASD-COMP vs COMP-COMP	Speech	64.706	72.727	68.717	0.673	0.000	*
ASD-COMP vs COMP-COMP	STACK	79.412	81.818	80.615	0.845	NaN	NA
BPD-COMP vs COMP-COMP	BODYsync	28.571	36.364	32.468	0.308	1.000	
BPD-COMP vs COMP-COMP	CROSSsync	50.000	61.364	55.682	0.594	0.282	
BPD-COMP vs COMP-COMP	CROSSturn	52.381	59.091	55.736	0.569	0.045	*
BPD-COMP vs COMP-COMP	FACEsync	64.286	63.636	63.961	0.670	0.007	*
BPD-COMP vs COMP-COMP	HEADsync	71.429	65.909	68.669	0.736	0.000	*
BPD-COMP vs COMP-COMP	INTRAsync	57.143	54.545	55.844	0.516	1.000	
BPD-COMP vs COMP-COMP	MovEx	69.048	70.454	69.751	0.686	0.242	
BPD-COMP vs COMP-COMP	Speech	59.524	63.636	61.580	0.693	0.007	*
BPD-COMP vs COMP-COMP	STACK	73.810	56.818	65.314	0.760	NaN	NA

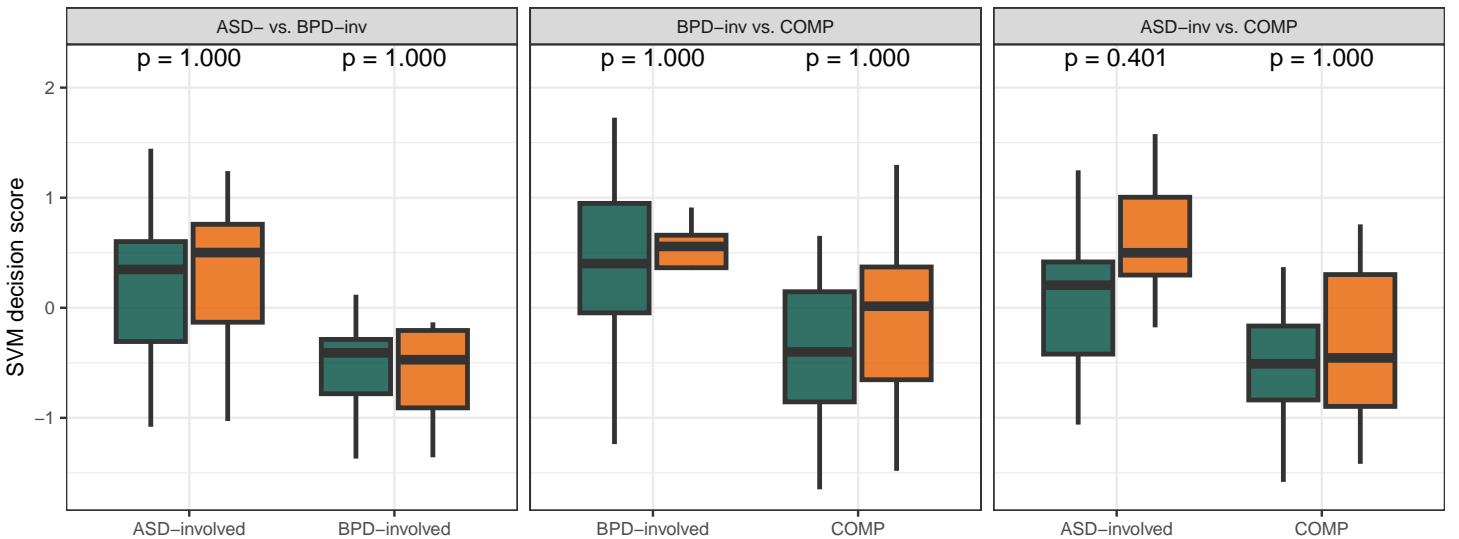
### 5.3 S5.3 Multi-group comparisons

comparison	model	BAC	p.fdr	sig
MultiGroup	BODYsync	48.6	1.000	
MultiGroup	CROSSsync	58.1	0.000	*
MultiGroup	CROSSturn	57.3	0.000	*
MultiGroup	FACEsync	57.6	0.007	*
MultiGroup	HEADsync	54.6	0.000	*
MultiGroup	INTRAsync	49.5	1.000	
MultiGroup	MovEx	60.8	0.000	*
MultiGroup	Speech	63.9	0.000	*
MultiGroup	STACK	63.2	NaN	NA

### 5.4 S5.4 Gender comparisons

Did the models perform better for one gender than the other? We perform unpaired Wilcoxon tests for the labels and models separately.

Gender comparison for decision scores

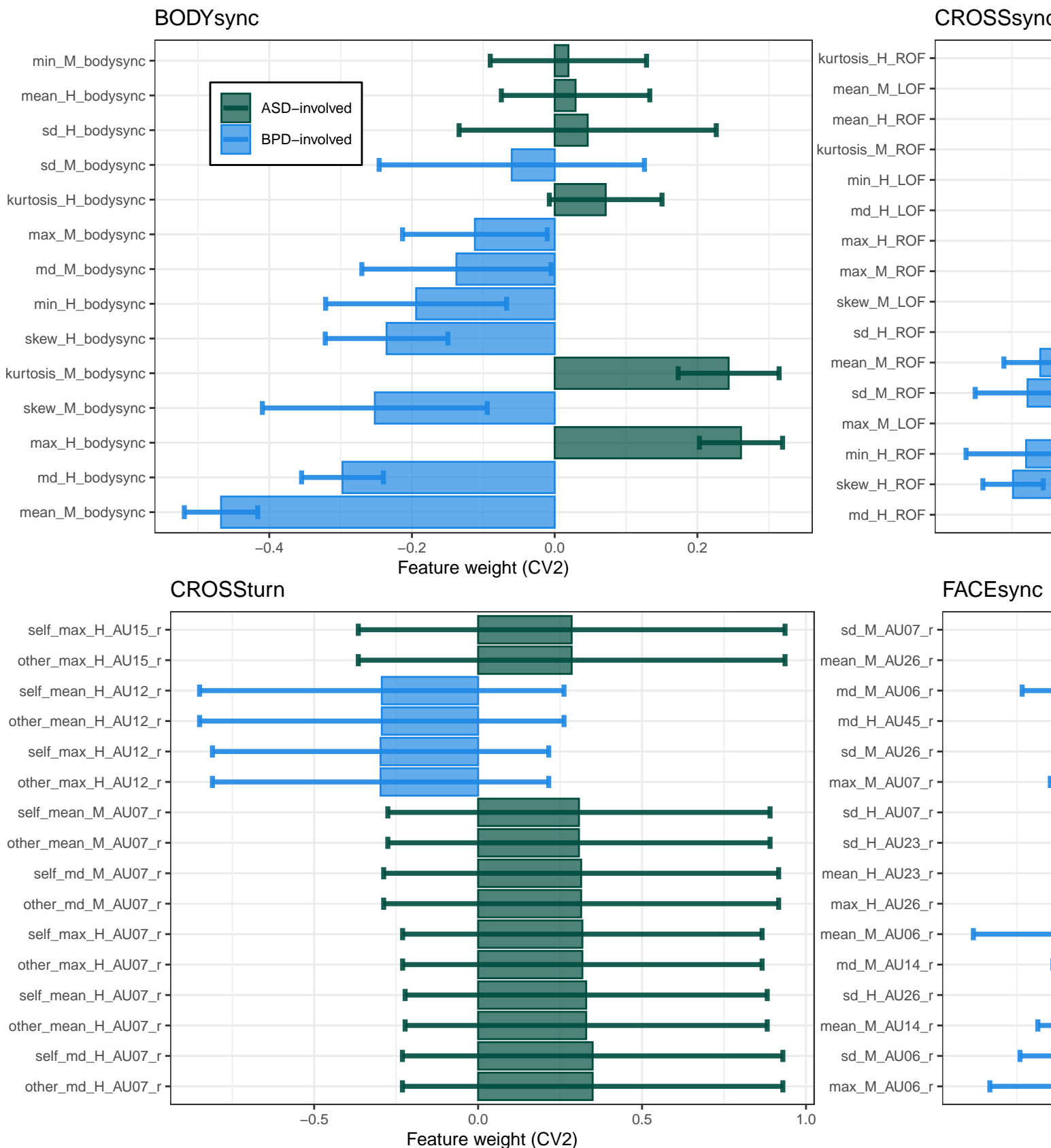


pwc: unpaired Wilcoxon test, FDR adjusted

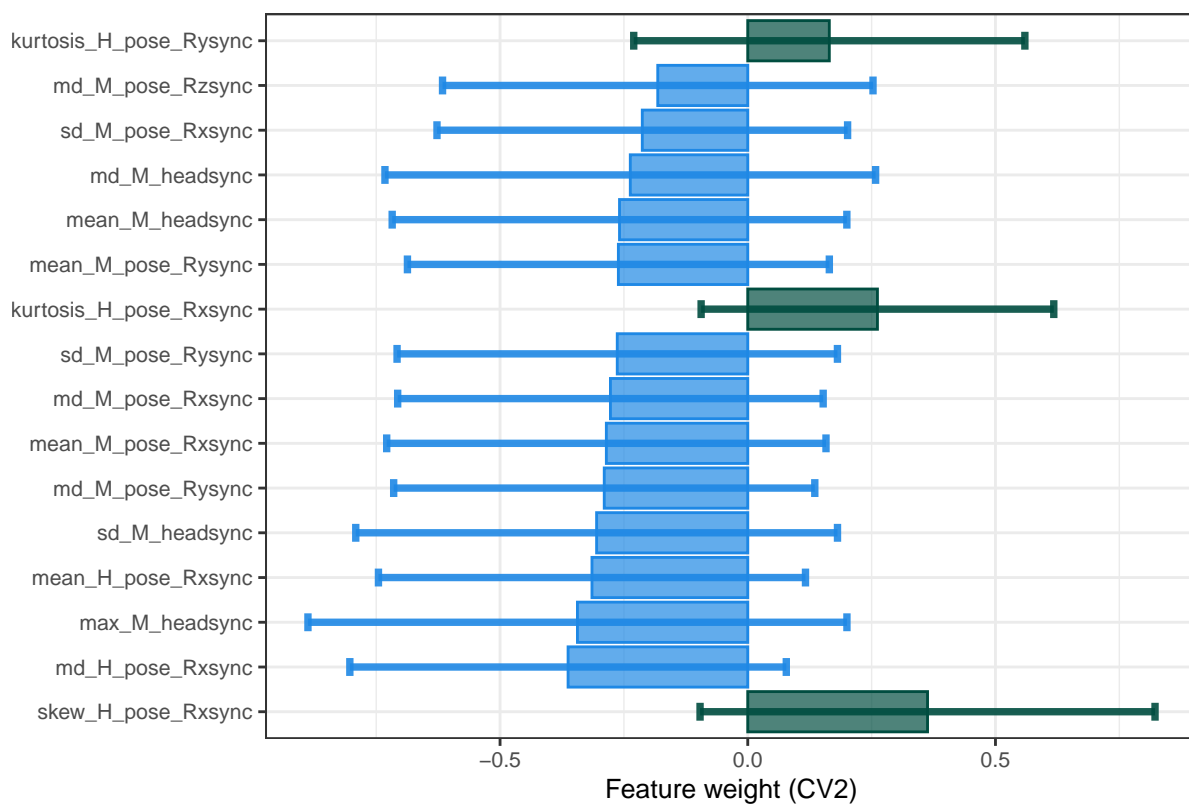
## 6 S6 Model visualisations

The following figures were created inspired by NeuroMiner<sup>15</sup> visualisations and show the sign-based consistency<sup>16</sup> as well as the feature weights of the models distinguishing between interaction partners from ASD- and BPD-involved interactions. For models with a large number of features, the top 16 features are plotted.

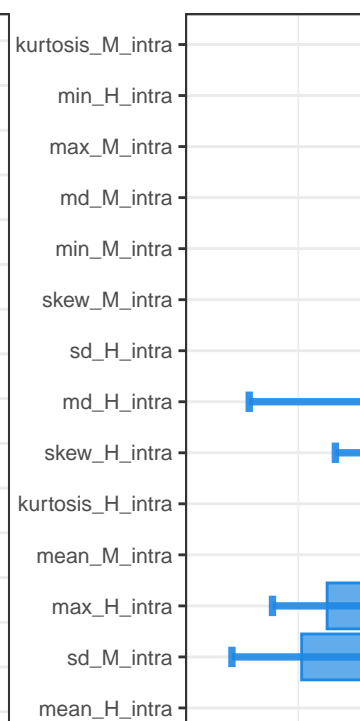
### 6.1 S6.1 ASD-involved versus BPD-involved classifiers



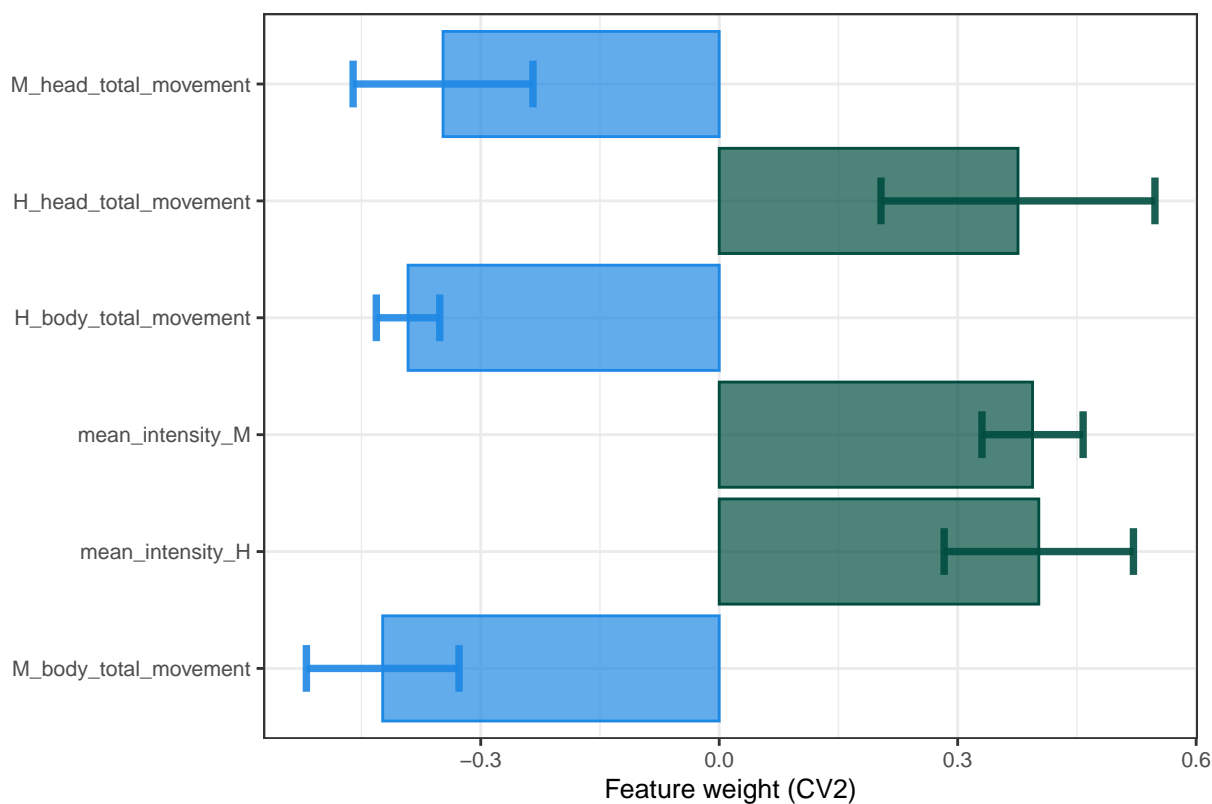
HEADsync



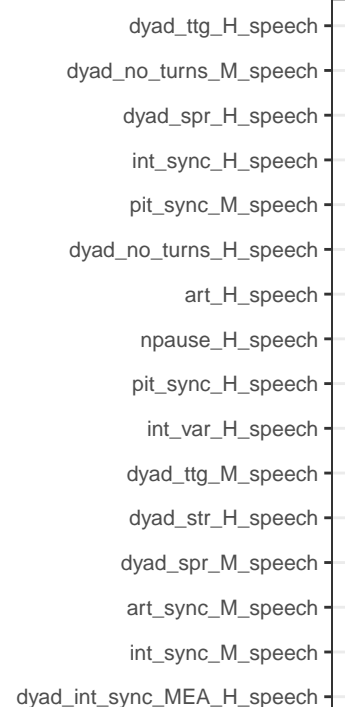
INTRAsync



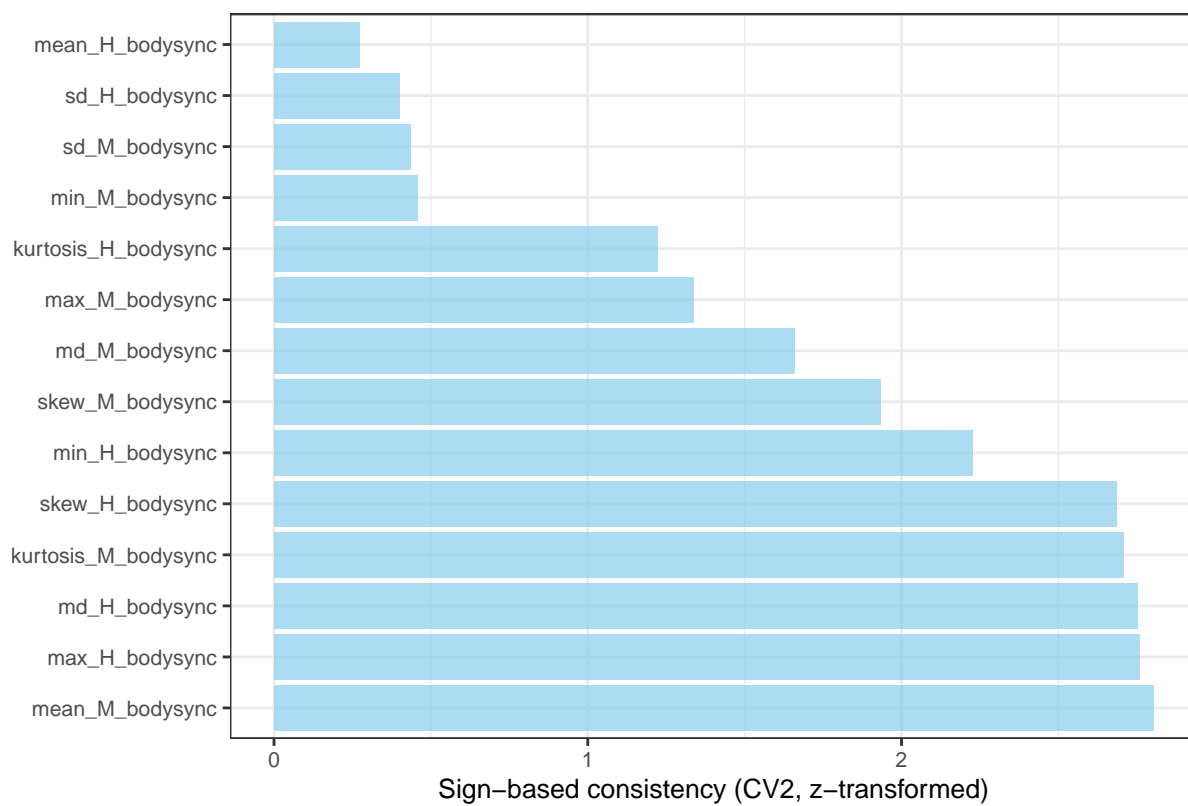
MovEx



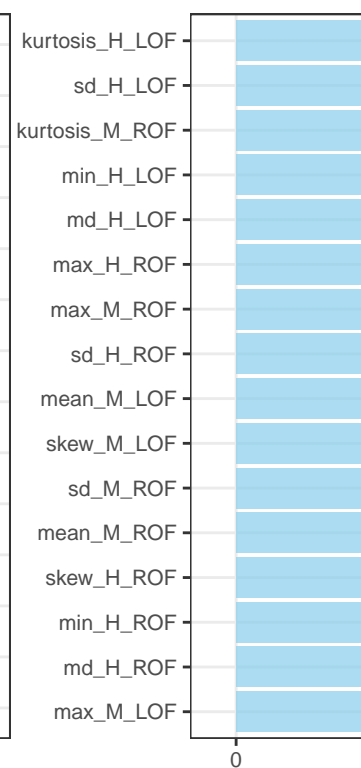
S



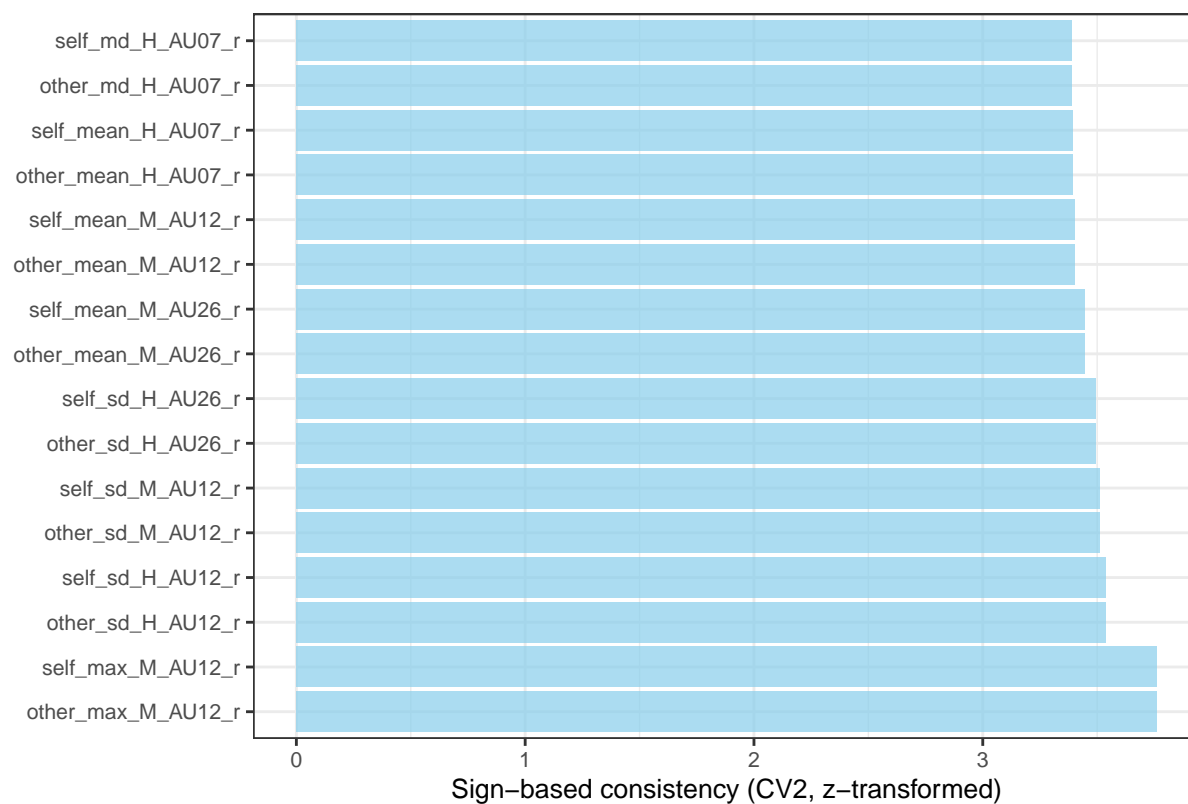
BODYsync



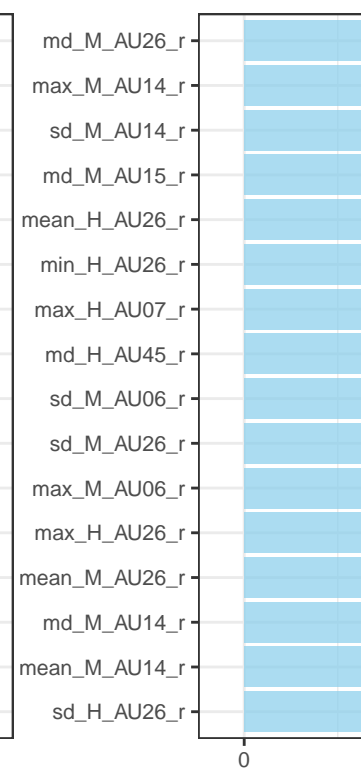
CROSSsync

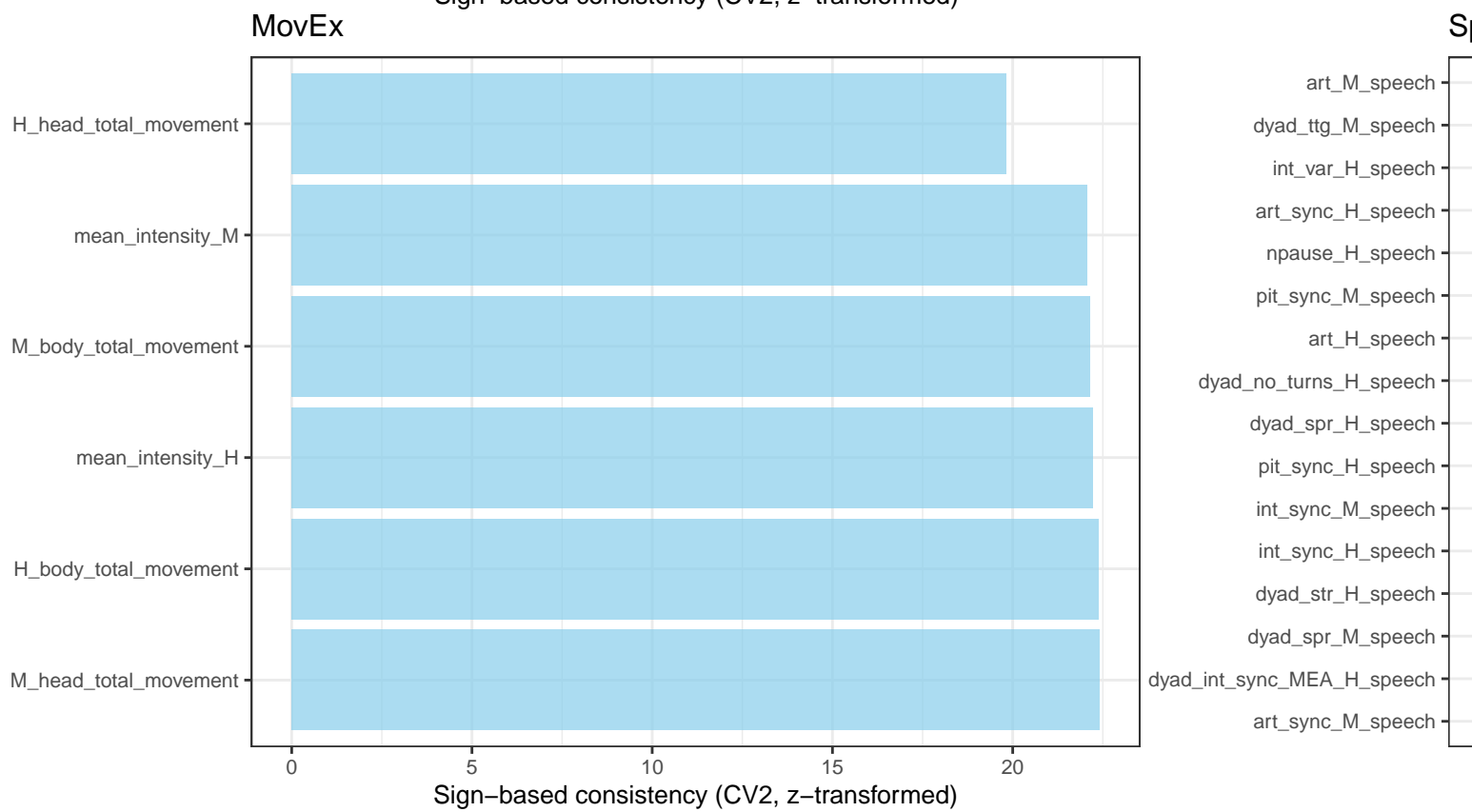
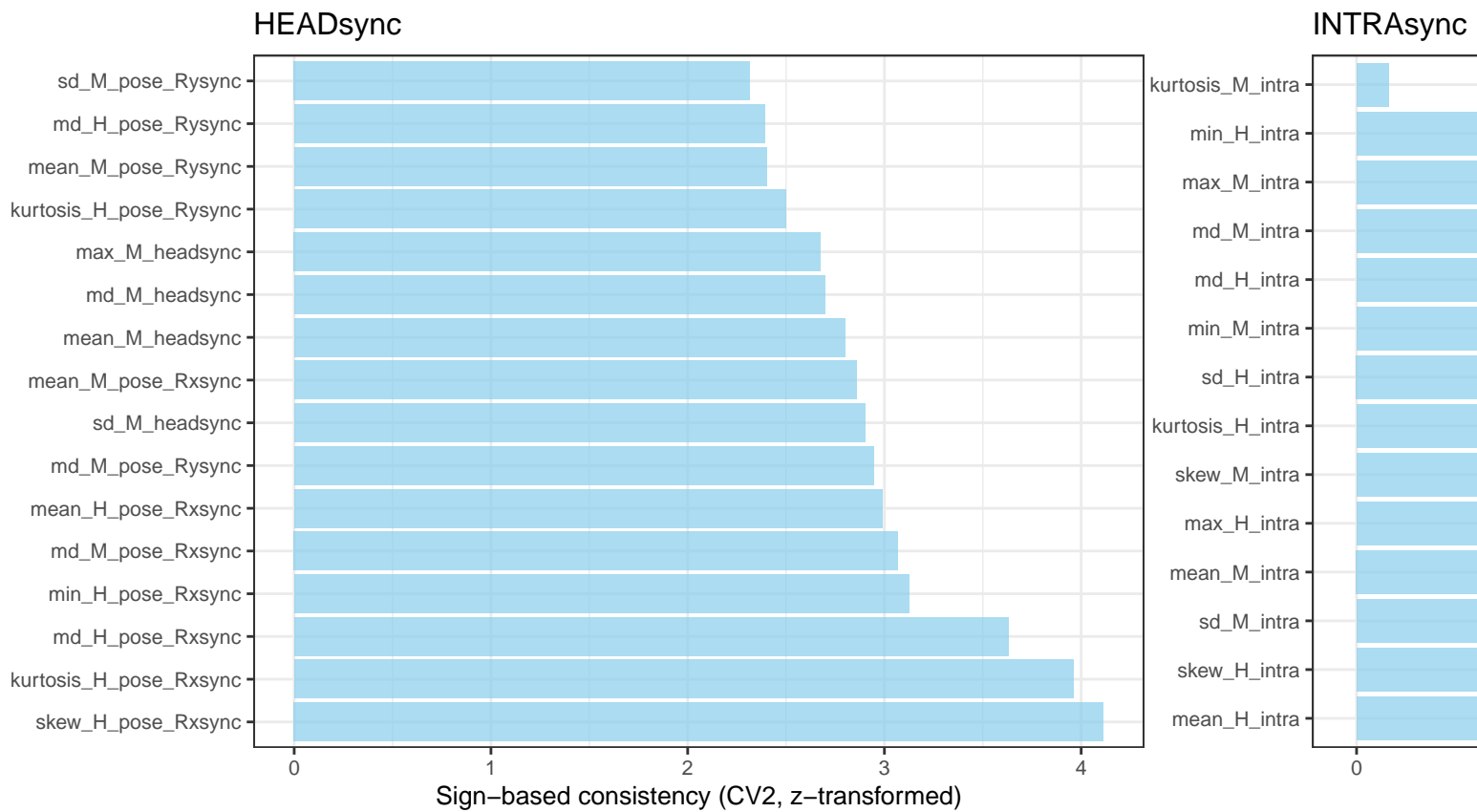


CROSSturn

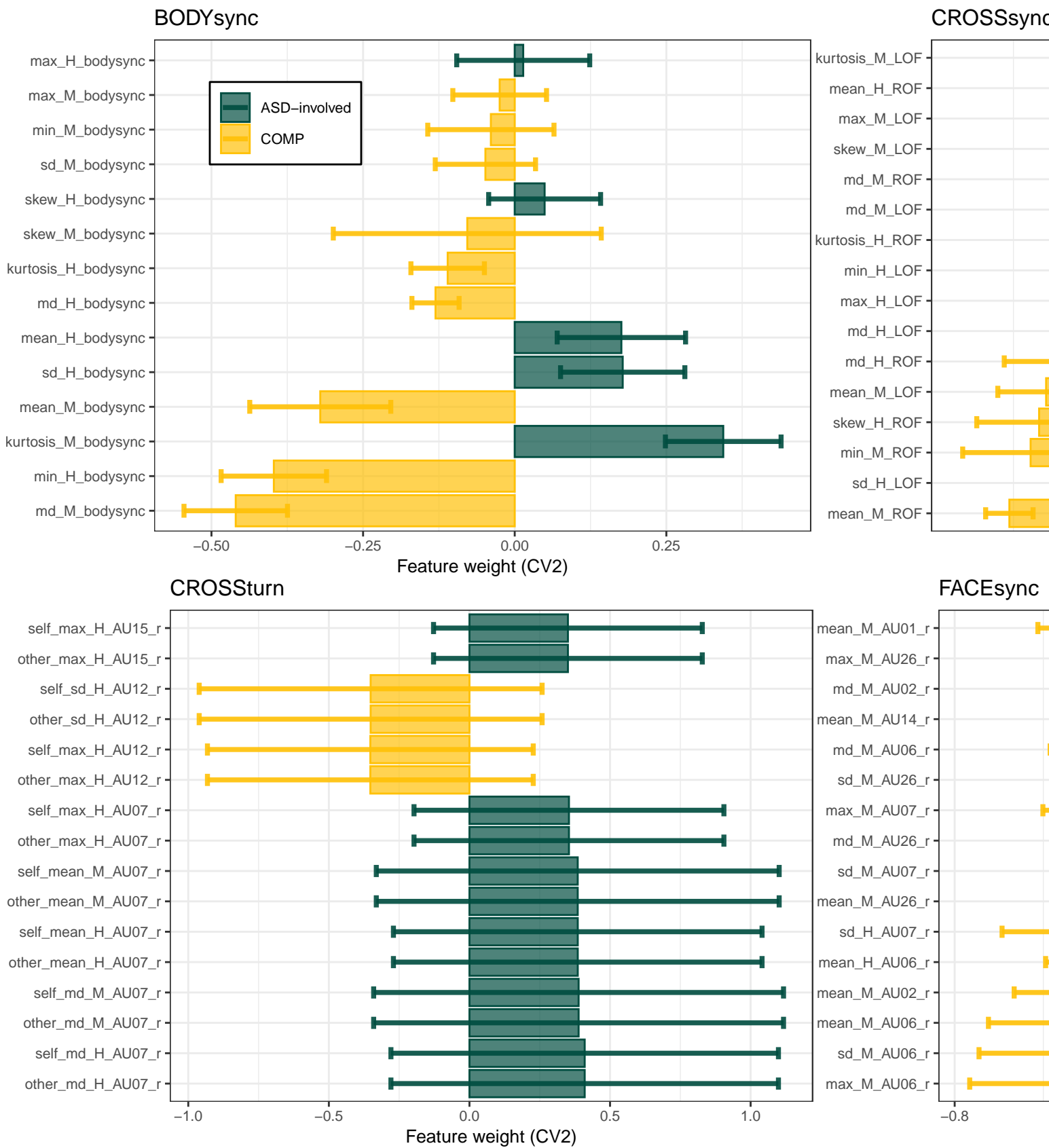


FACEsync

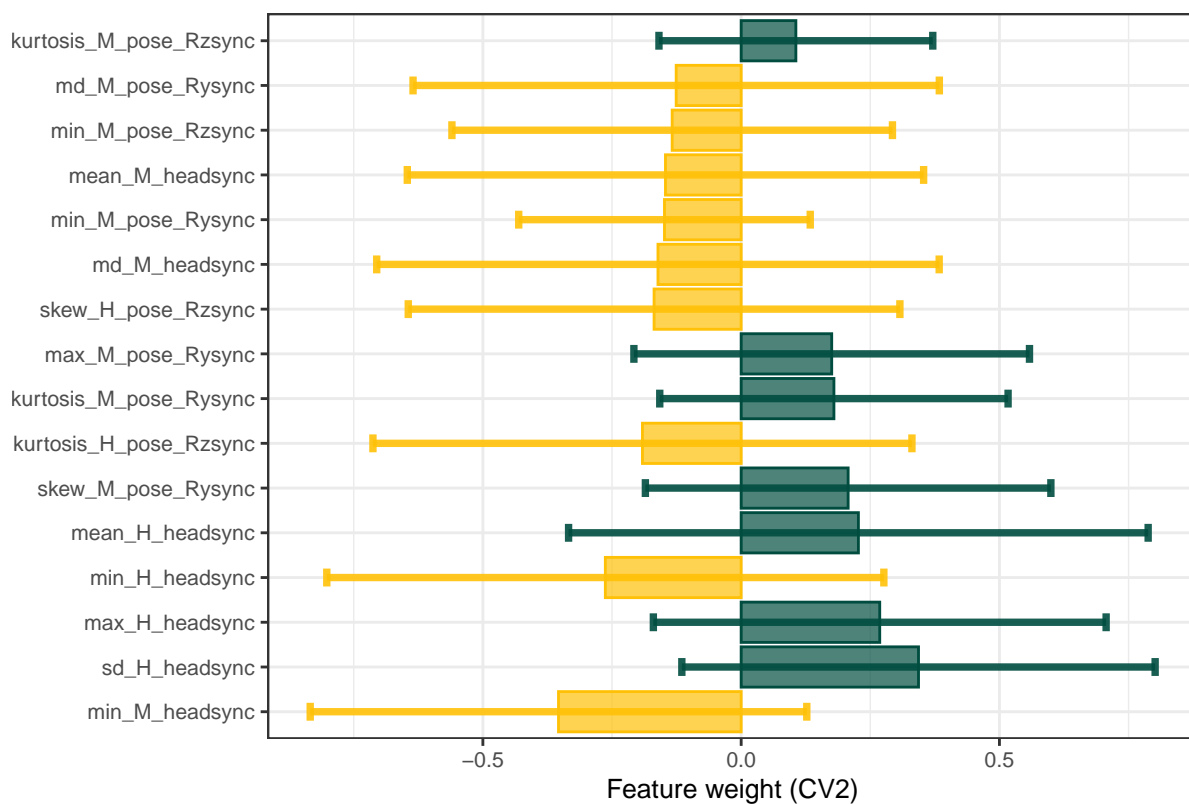




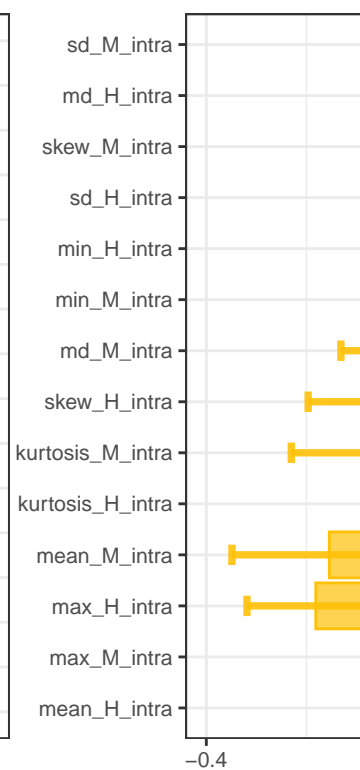
## 6.2 S6.2 ASD-involved versus COMP classifiers



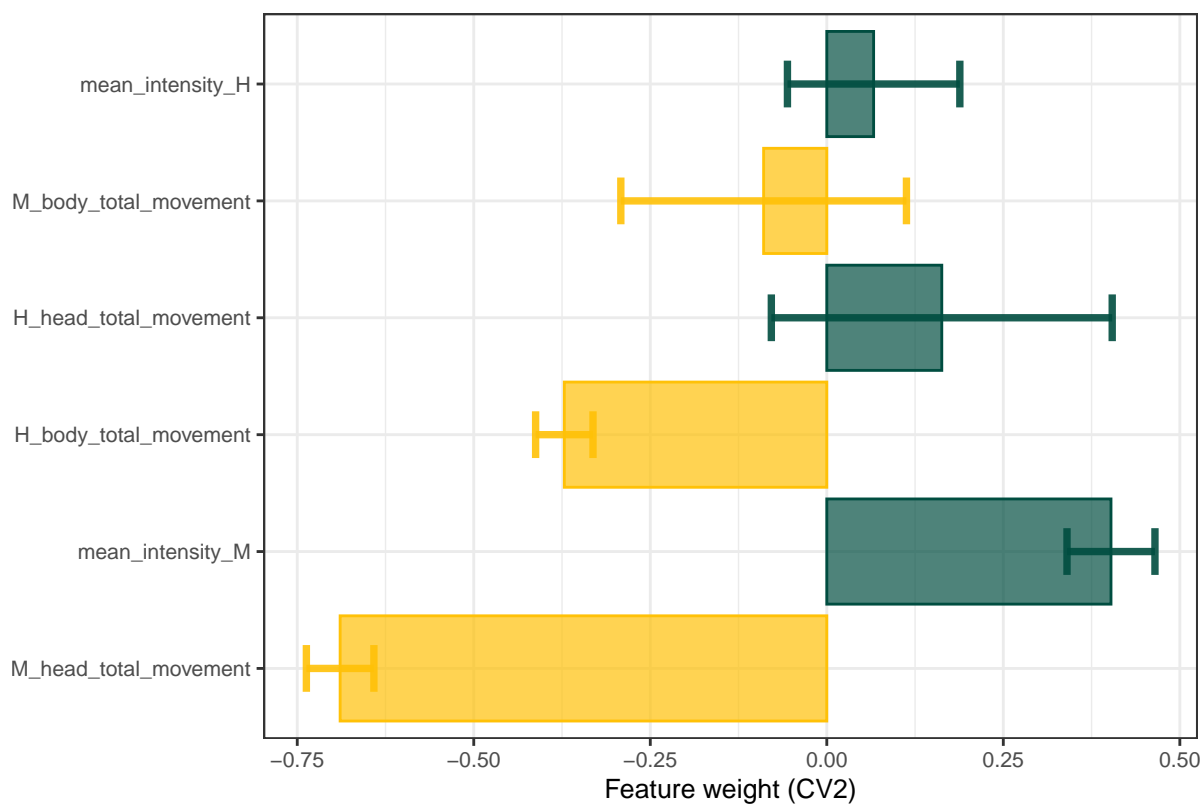
HEADsync



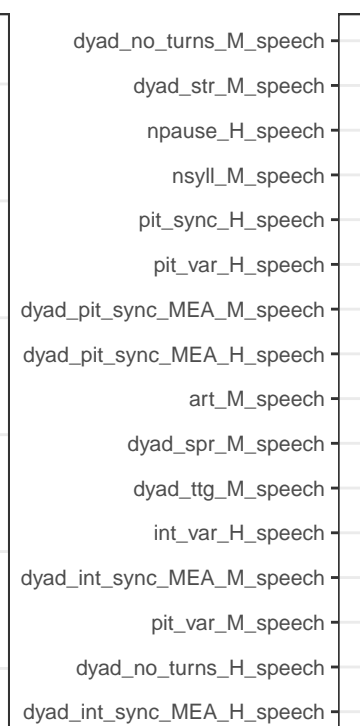
INTRAsync

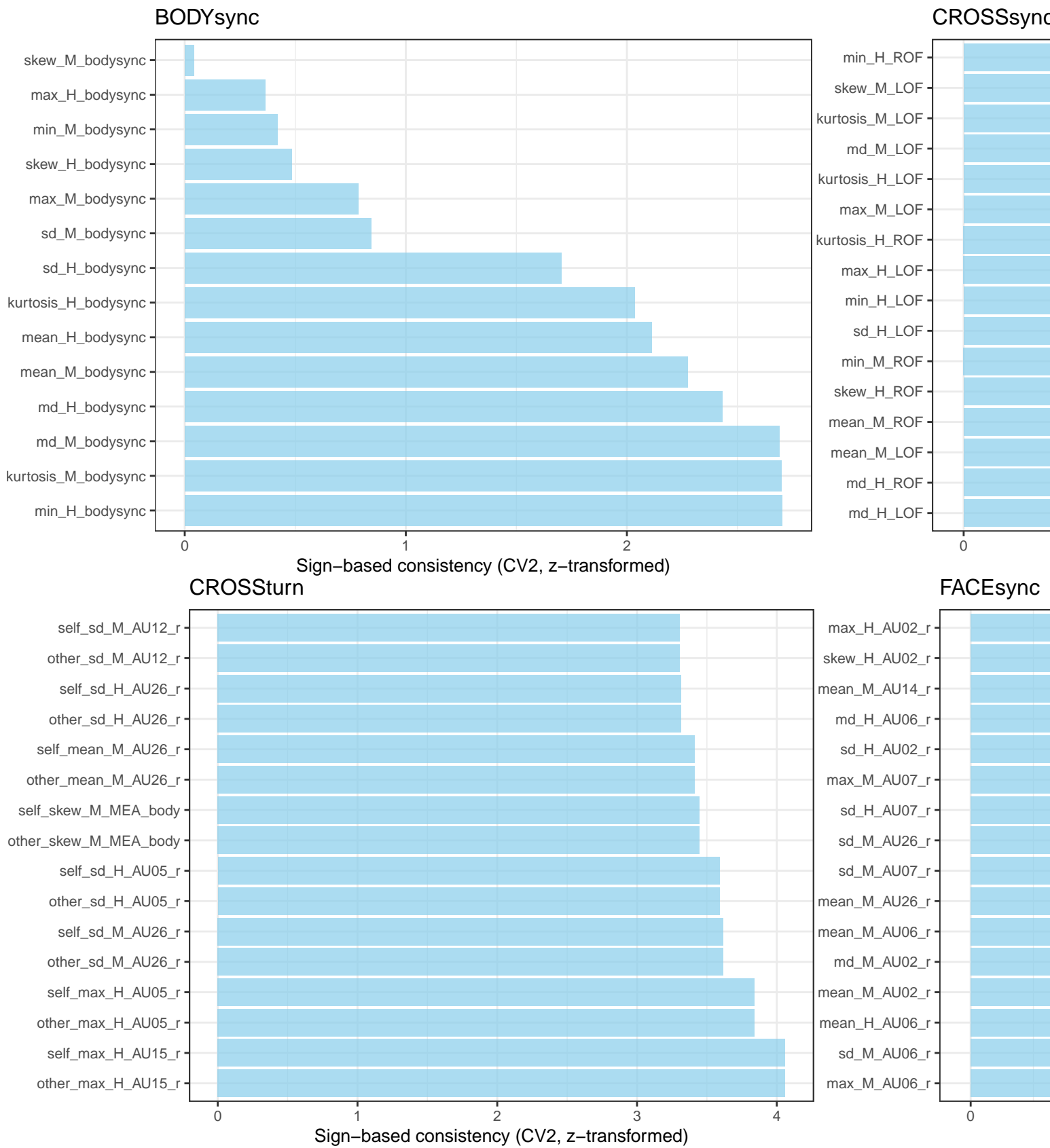


MovEx



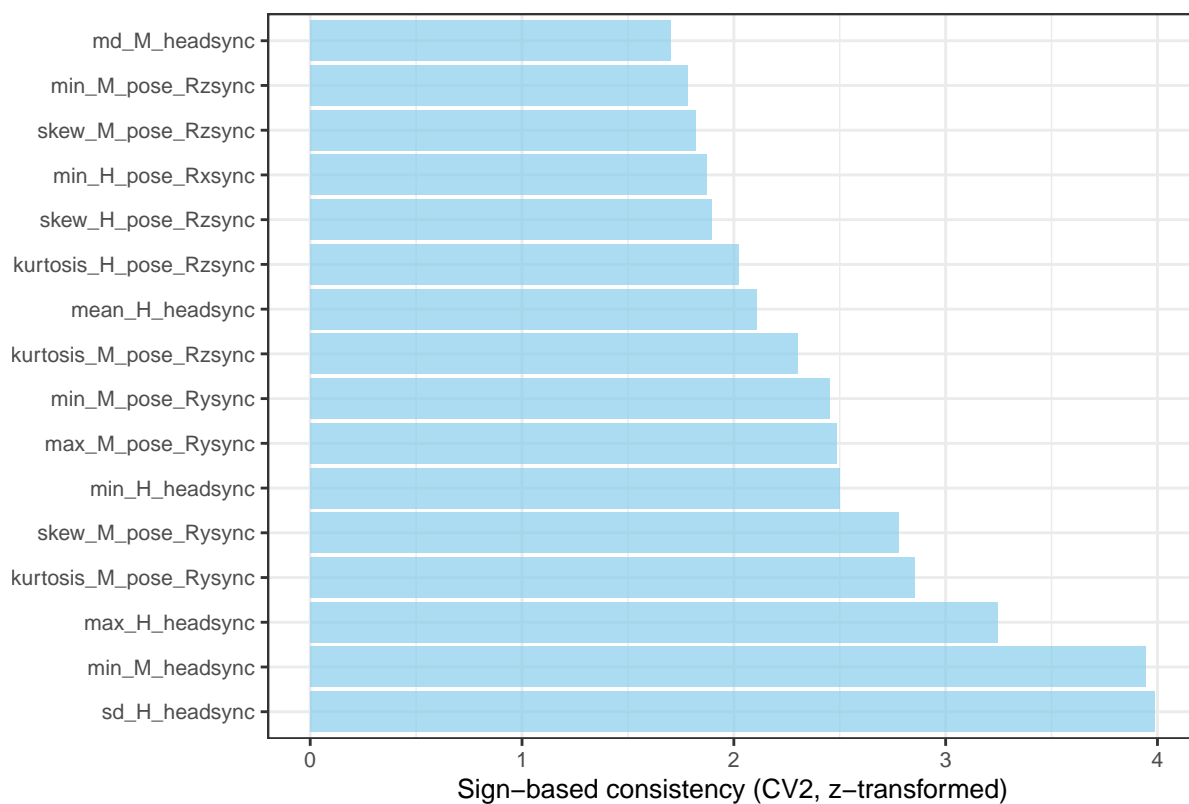
S



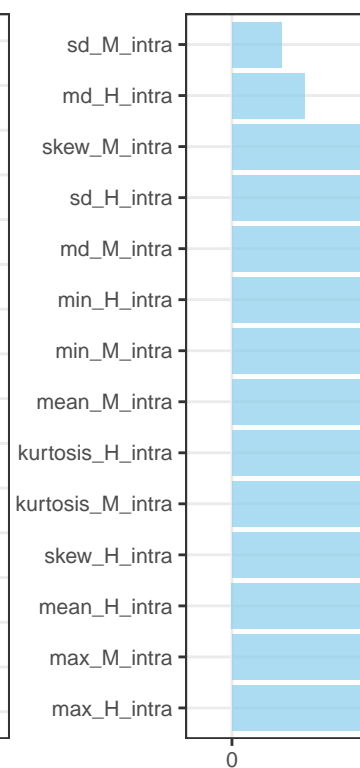




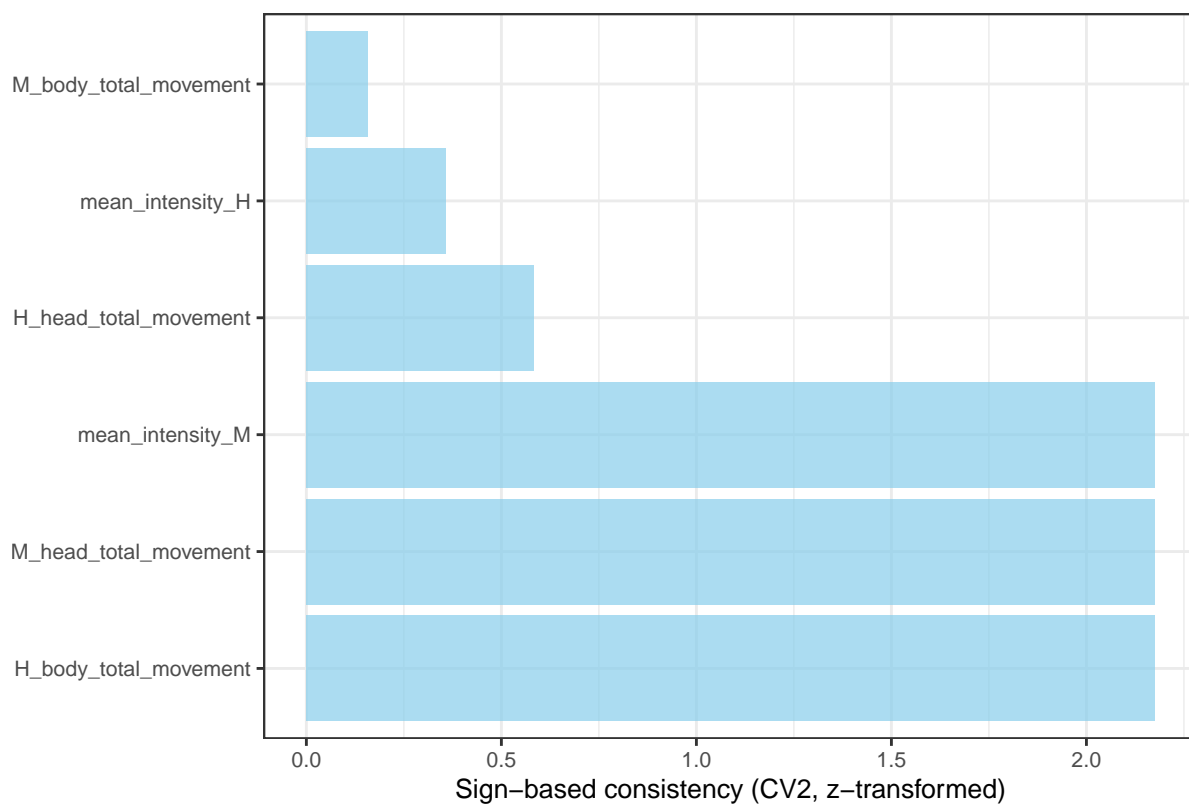
HEADsync



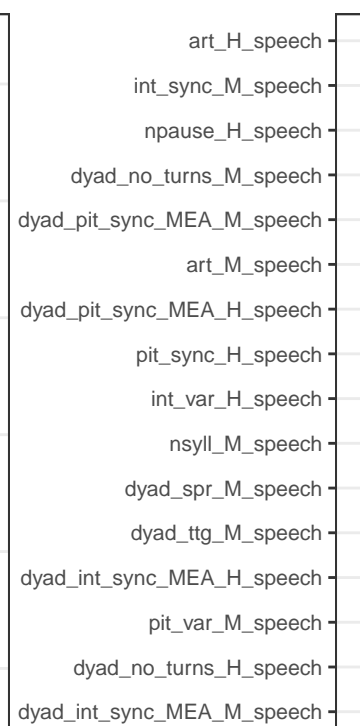
INTRAsync



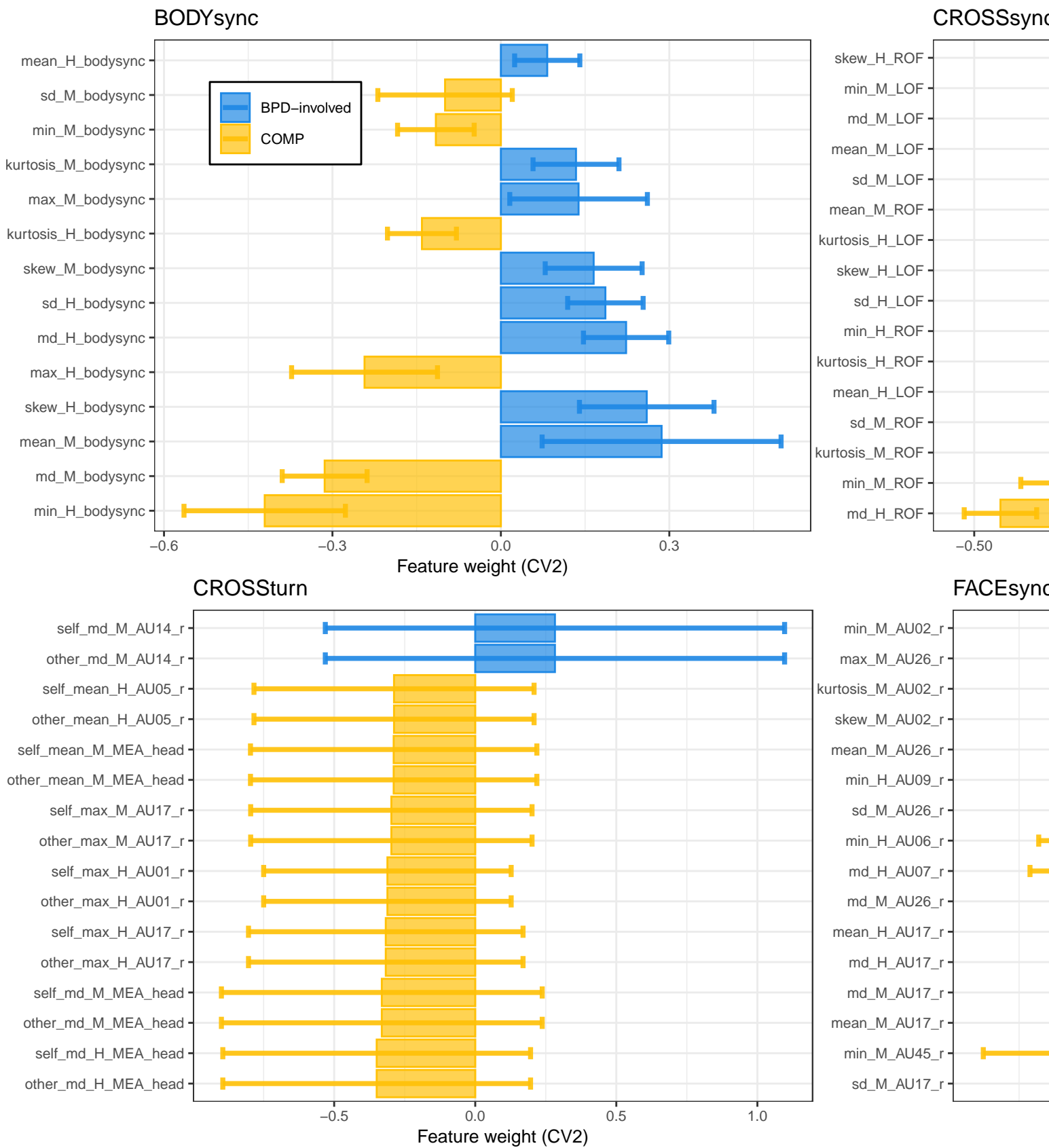
MovEx

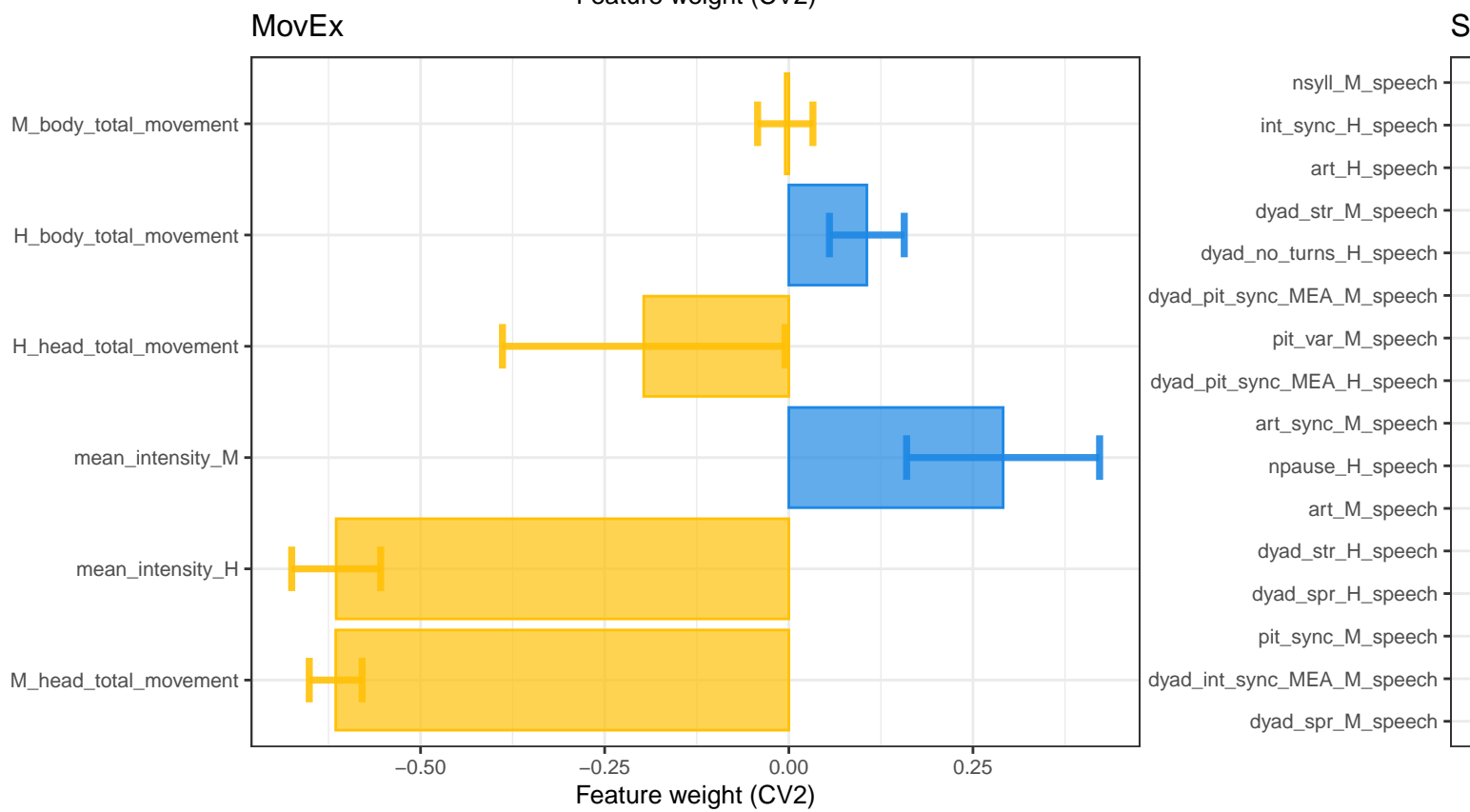
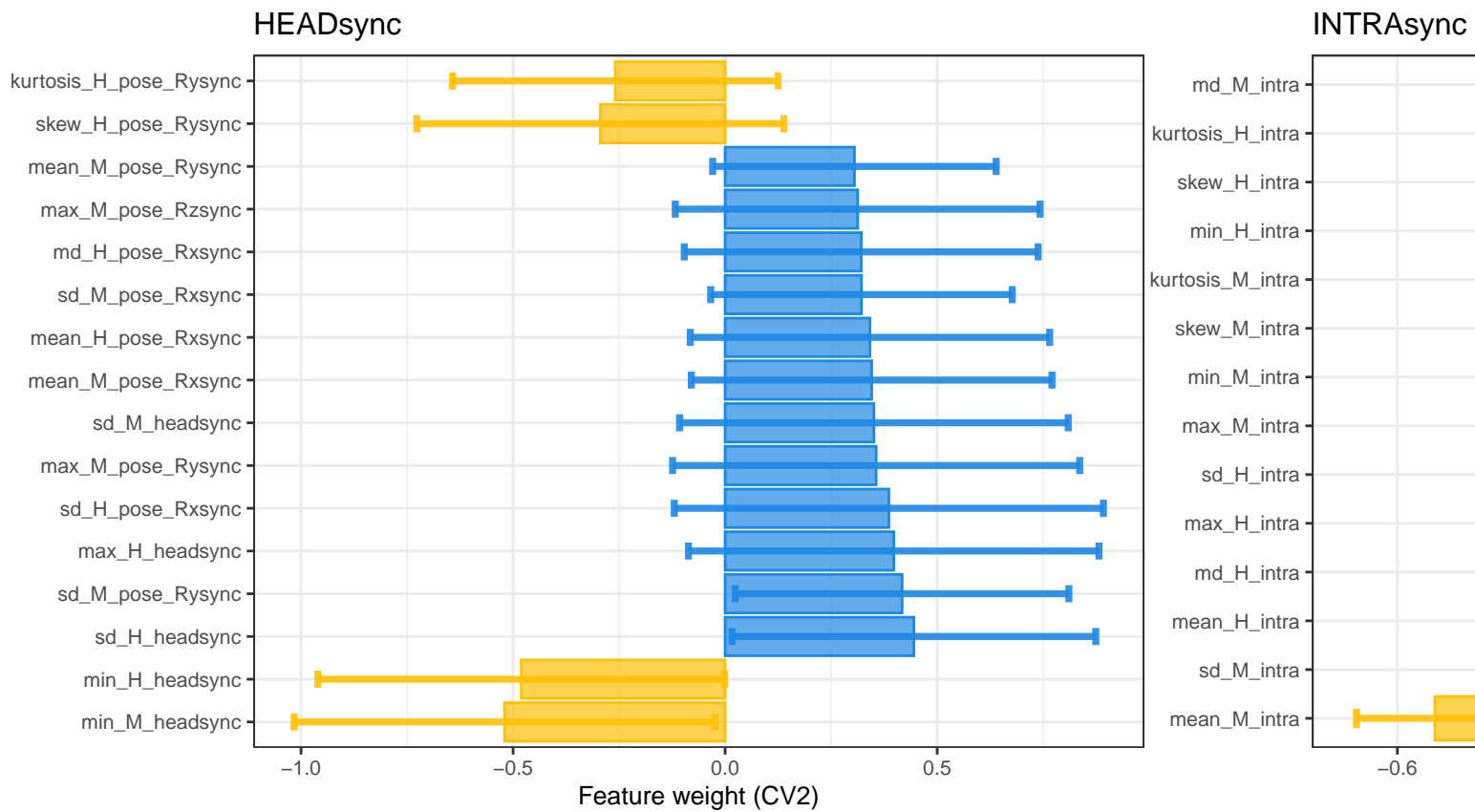


S

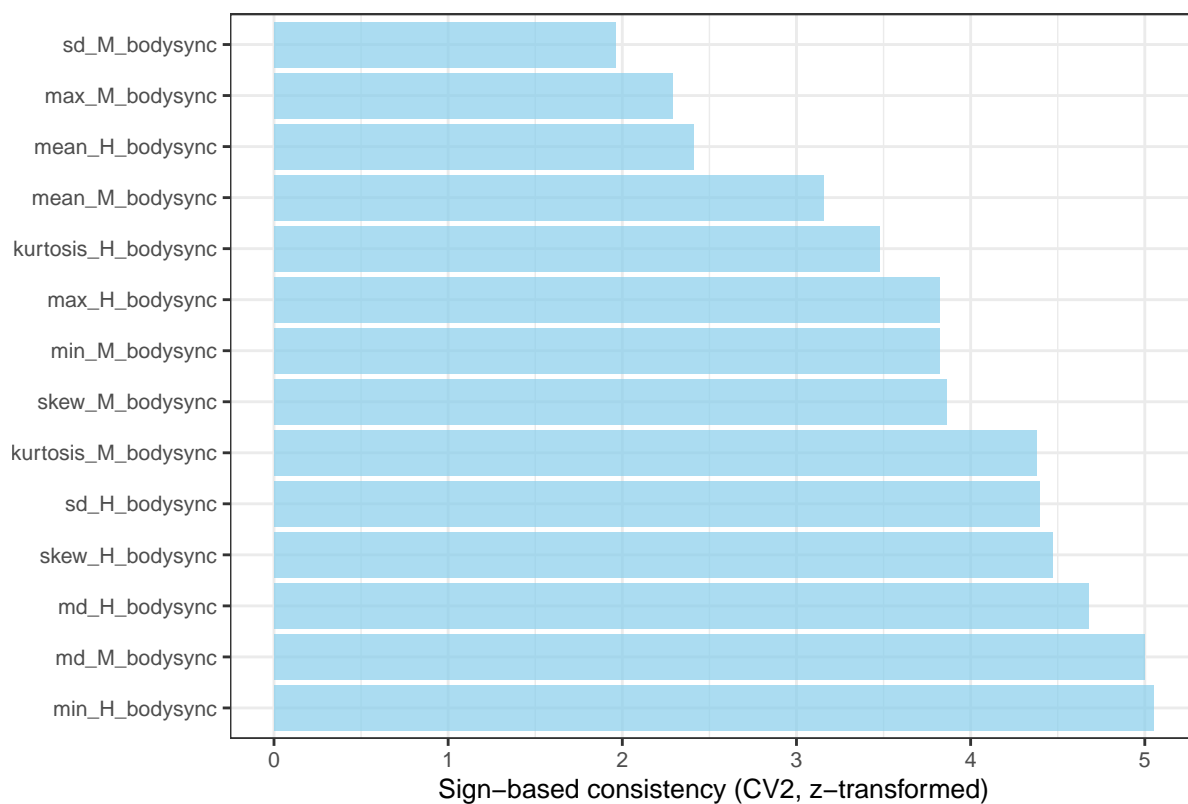


### 6.3 S6.3 ASD-involved versus COMP classifiers

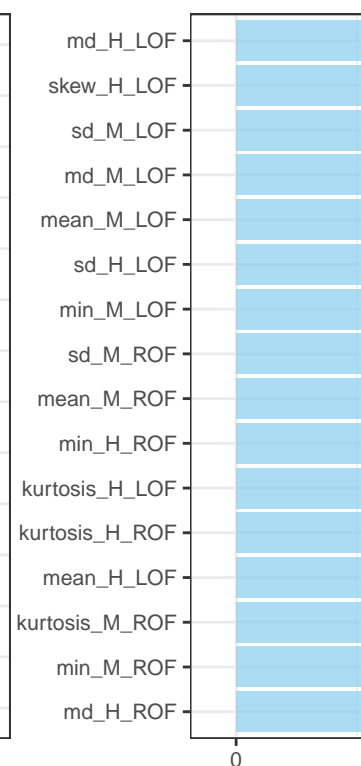




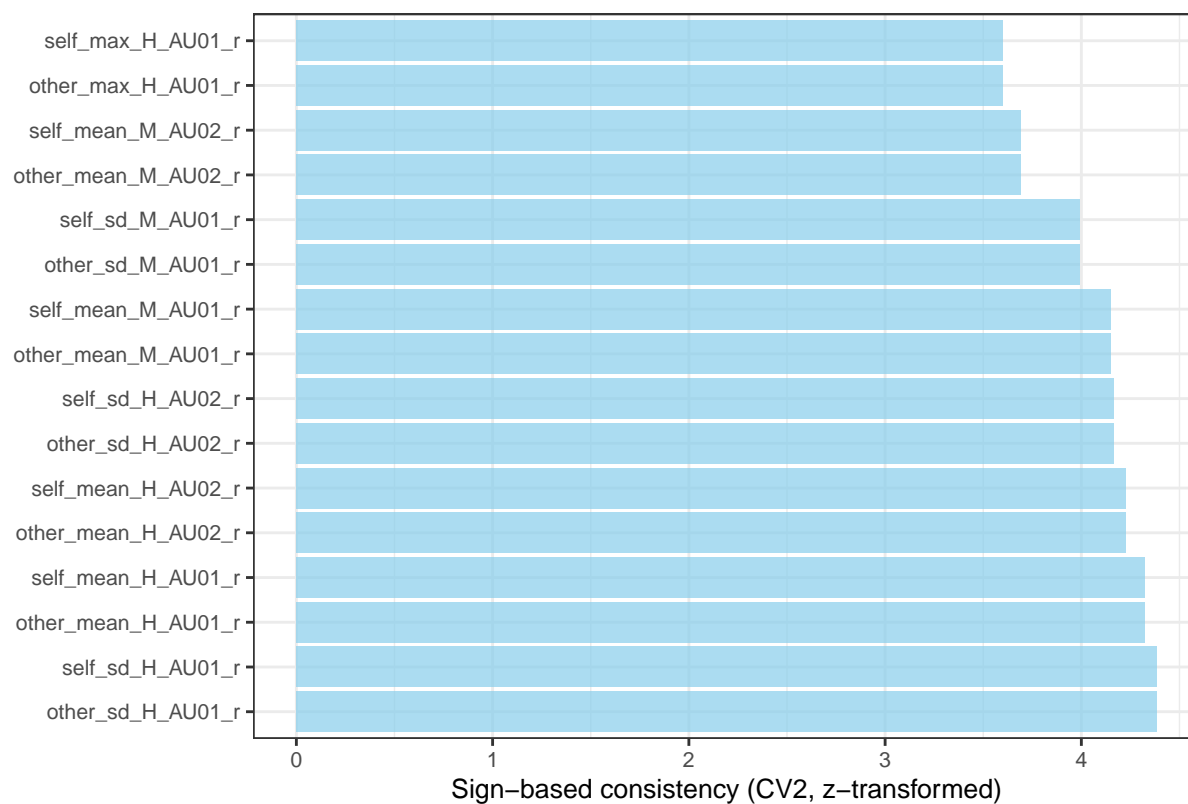
BODYsync



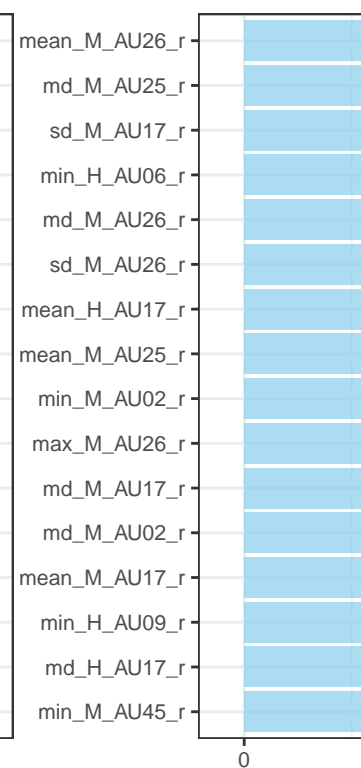
CROSSsync

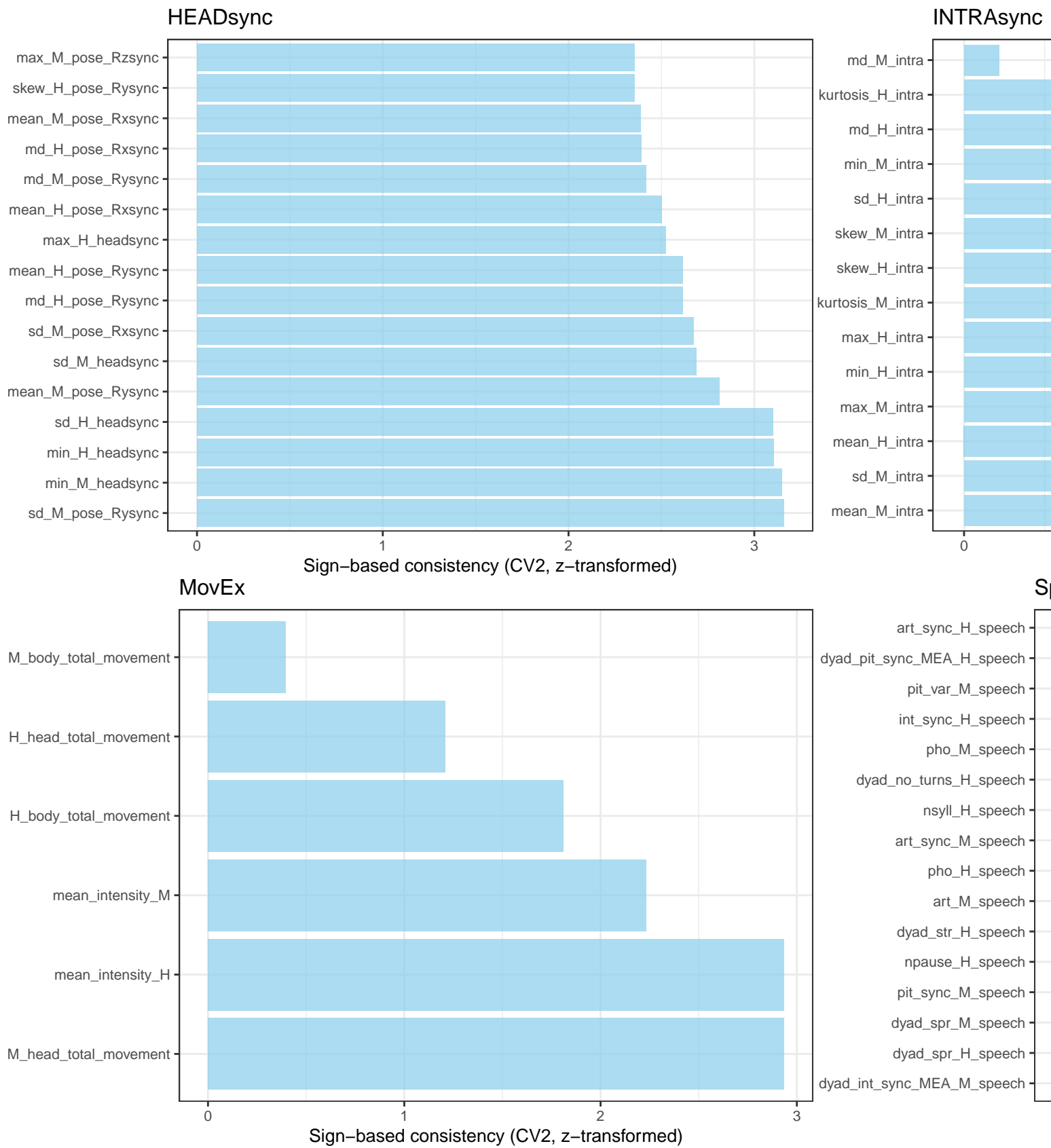


CROSSsturn



FACEsync





## 7 S7 References

1. Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J. & Clubley, E. The Autism-Spectrum Quotient (AQ): Evidence from Asperger Syndrome/High-Functioning Autism, Males and Females, Scientists and Mathematicians. *Journal of Autism and Developmental Disorders* **31**, 5–17 (2001).

2. Paulus, C. Der Saarbrücker Persönlichkeitsfragebogen SPF (IRI) zur Messung von Empathie: Psychometrische Evaluation der deutschen Version des Interpersonal Reactivity Index. (2009).
3. Popp, K. *et al.* Faktorstruktur and Reliabilität der Toronto-Alexithymie-Skala (TAS-20) in der deutschen Bevölkerung. *Psychotherapie · Psychosomatik · Medizinische Psychologie* **58**, 208–214 (2008).
4. Hautzinger, M., Bailer, M., Worall, H. & Keller, F. Beck-depressions-inventar (BDI). (1994).
5. Graf, A. A German version of the Self-Monitoring Scale. *Zeitschrift für Arbeits- und Organisationspsychologie* **48**, 109–121 (2004).
6. Kirby, A., Edwards, L., Sugden, D. & Rosenblum, S. The development and standardization of the Adult Developmental Co-ordination Disorders/Dyspraxia Checklist (ADC). *Research in Developmental Disabilities* **31**, 131–139 (2010).
7. Bohus, M. *et al.* The Short Version of the Borderline Symptom List (BSL-23): Development and Initial Data on Psychometric Properties. *Psychopathology* **42**, 32–39 (2008).
8. Costa, A. P., Steffgen, G. & Samson, A. C. Expressive Incoherence and Alexithymia in Autism Spectrum Disorder. *J Autism Dev Disord* **47**, 1659–1672 (2017).
9. Koehler, J. C. *et al.* Machine learning classification of autism spectrum disorder based on reciprocity in naturalistic social interactions. *Transl Psychiatry* **14**, 1–9 (2024).
10. Boersma, P. & Weenink, D. Praat: Doing phonetics by computer. (2022).
11. Hirst, D. The analysis by synthesis of speech melody: From data to models. *Journal of Speech Sciences* **1**, 55–83 (2011).
12. Plank, I. S., Koehler, J. C., Nelson, A., Koutsouleris, N. & Falter-Wagner, C. Automated extraction of speech and turn-taking parameters in autism allows for diagnostic classification using a multivariable prediction model. *Frontiers in Psychiatry* **14**, 1257569 (2023).
13. De Jong, N. H., Pacilly, J. & Heeren, W. Uhm-o-meter [Computer software]. (2021).
14. De Jong, N. H., Pacilly, J. & Heeren, W. PRAAT scripts to measure speed fluency and breakdown fluency in speech automatically. *Assessment in Education: Principles, Policy and Practice* **28**, 456–476 (2021).
15. Koutsouleris, N., Vetter, C., Wiegand, A., Hahn, L. & Mena, S. Neurominer. (2024).
16. Gómez-Verdejo, V., Parrado-Hernández, E., Tohka, J. & Initiative, A. D. N. Sign-consistency based variable importance for machine learning in brain imaging. *Neuroinformatics* **17**, 593–609 (2019).