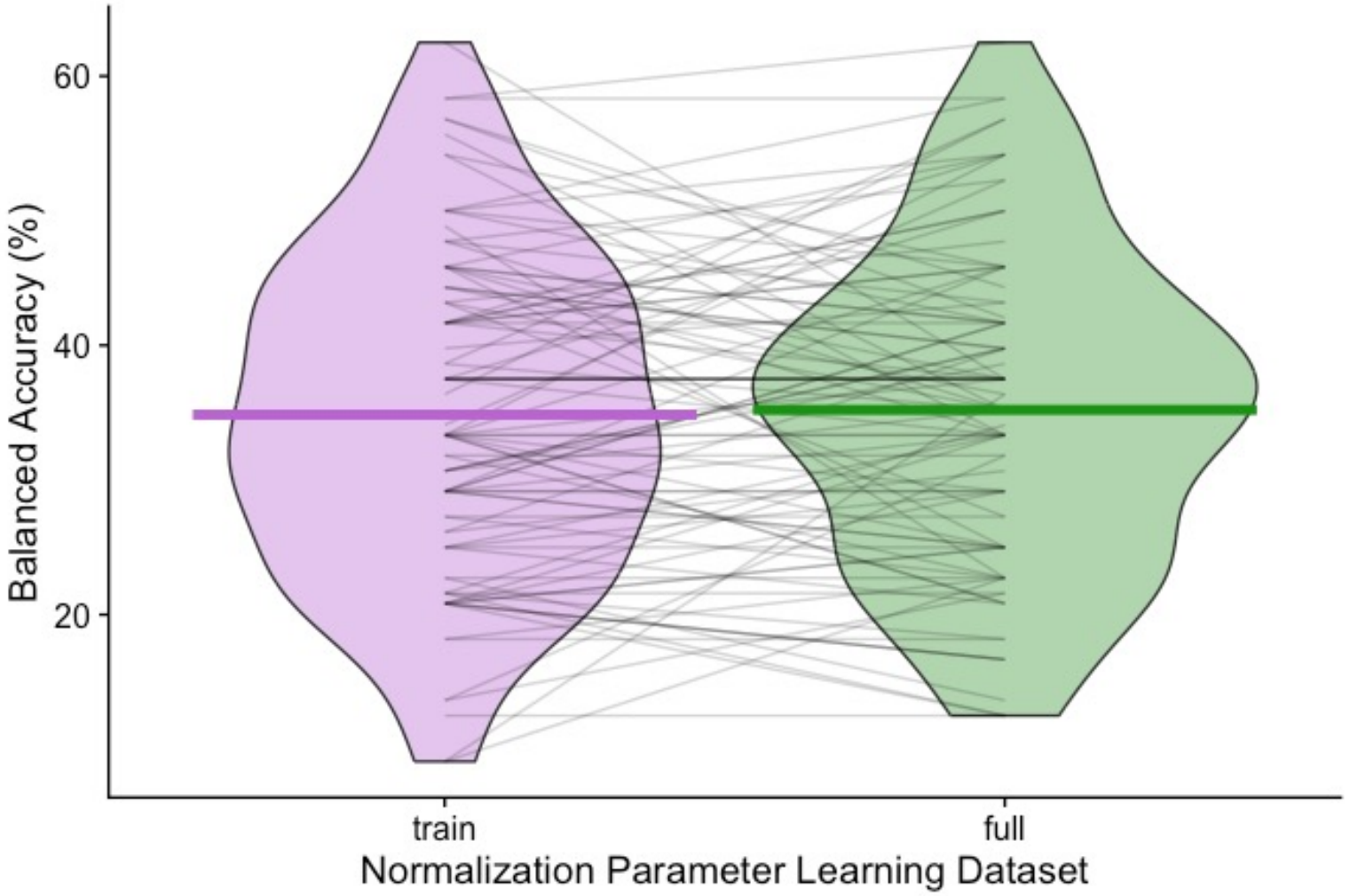
Effect of normalization on **train and test folds together** vs. based on **training data only**

For **low_freq_power** in **ADHD**, SVMs were run using a **linear kernel** with the scikit-learn **SVC()** function, with 10-fold CV repeated 10 times for each iteration. Each model was normalized using either the median and IQR from the training set only (**purple**) or from the full input matrix (**green**), with the latter allowing for potential data leakage for the unseen test set. While there is variability in individual test folds (shown as black lines going across the two violins), the mean balanced accuracy is very similar for the two model types: **34.8% (train)** vs **35.2% (full)**



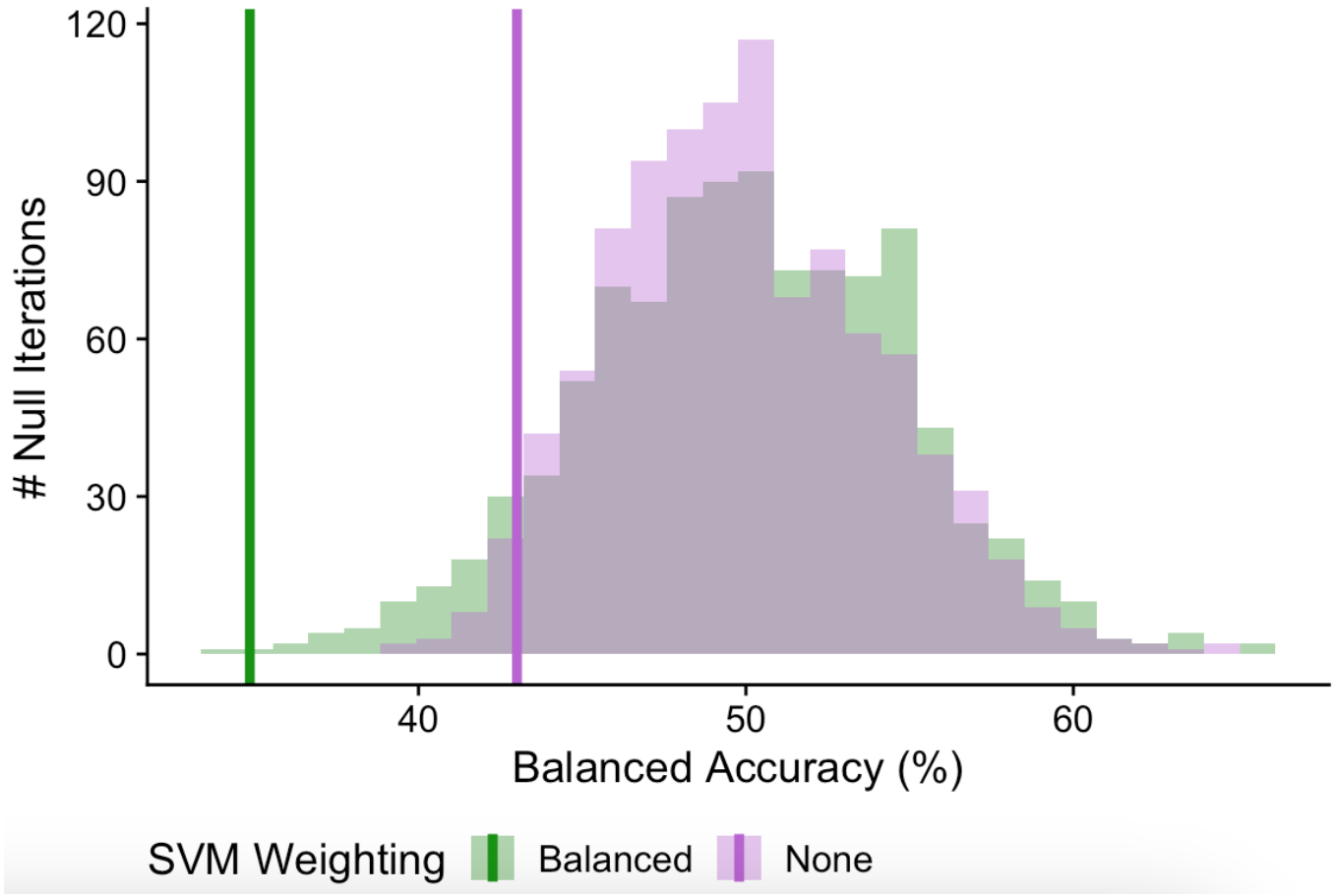
Null distributions **with** vs. **without** reweighting

Here, null permutations were run 1000 times using the same classification code as the main analyses.

Null distributions are shown as histograms indicating whether inverse probability weighting was applied (**green**) or no reweighting was applied (**purple**).

Main balanced accuracy values corresponding to each weighting type are shown as vertical lines.

While the balanced accuracy is low regardless of SVM weighting type, it is higher with **no reweighting** (**43.0%**) versus with **inverse probability reweighting** (**34.8%**).

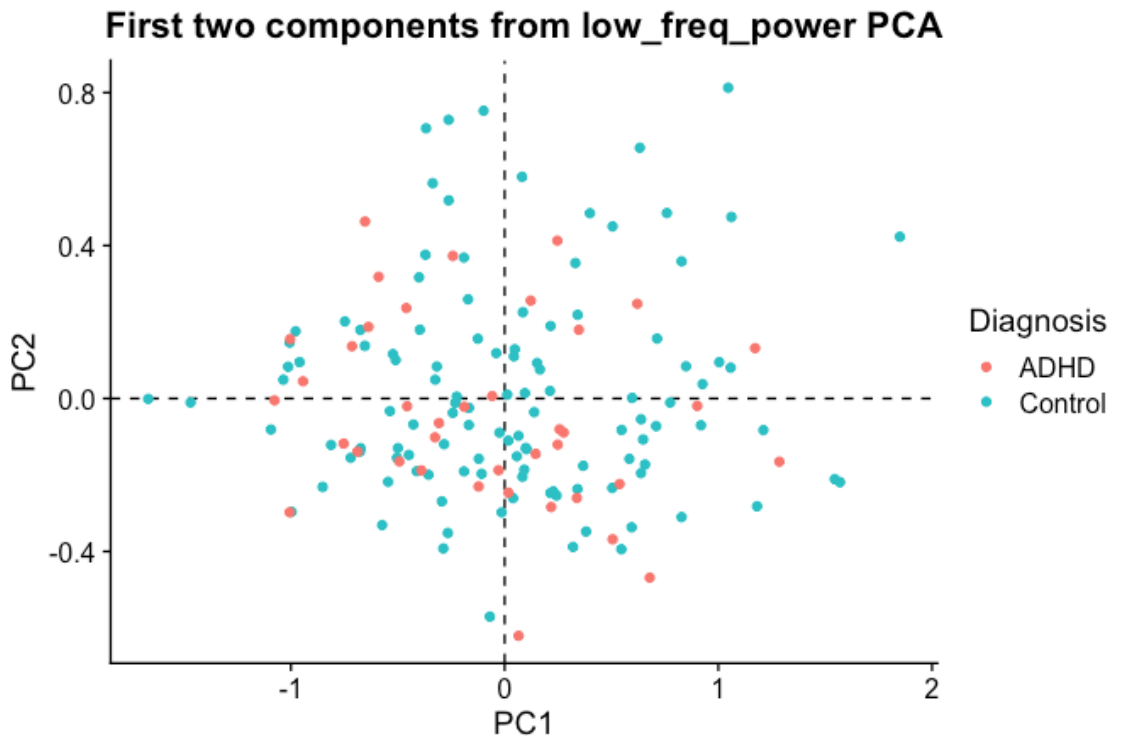


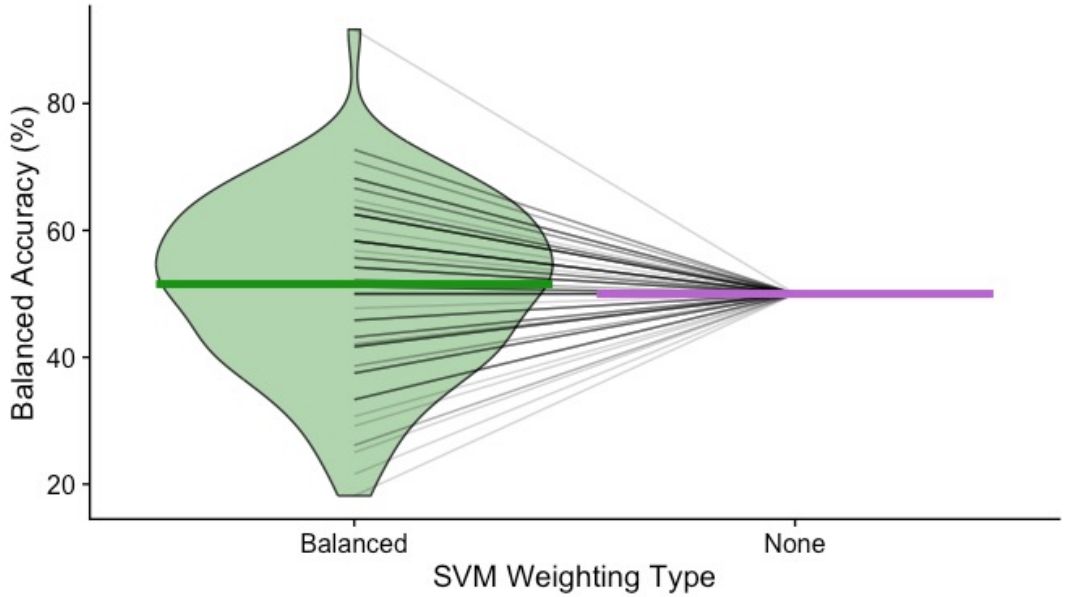
Using the first two principal components for classification

Here, we performed a principal components analysis (PCA) on the low_freq_power data for all ADHD and control participants, after standardizing the values per brain region using the mixed sigmoid normalization method.

We retained the first two components, PC1 and PC2, for which individual scores are shown in the top right plot. There isn't much visible distinction between **ADHD** and **control** participants in this space.

We then used these two PCs as the input feature matrix for a linear SVM using either **inverse probability weighting (green)** or **no reweighting (purple)**, and found both yielded means around 50% -- though there was a wide distribution only when using inverse probability weighting.





Effects of (A) class weighting and (B) SVM kernel on low_freq_power performance in ADHD

Here, each classifier kernel type was run once with **inverse probability weighting** (using the class_weight='balanced' option), as shown in teal and pink; or with **no reweighting**, as shown in purple and green.

This reveals that the balanced accuracy is much more negatively skewed with the linear kernel, particularly with inverse probability weighting.

For the RBF kernel, the balanced accuracy varies with a median of 50% with inverse probability weighting; without this weighting, all iterations yielded balanced accuracy of 50%, corresponding to the classifier predicting "Control" every time as the majority class.