**Standard Preprocessing Steps**

**Developmental Neuroscience Lab**

**Child and Youth Studies**

All of these steps will be done on each participant’s data individually.

**Sanity Check**

Before preprocessing, do a quick check to make sure all the data is in its place. First, check the primary window (left side) and make sure each participant’s data has been properly loaded. Second, scroll through all the channels to make sure each channel is present (yellow arrows on the top bar). Third, check that the length of the continuously recorded data makes senses (time is found at the bottom; length can be checked by scrolling through the data and looking at both the start time and end time). Finally, check that the markers are in place (intermittent codes along the bottom of the data).

**1. Edit Channels**

Remove any channels that did not record properly for a substantial portion of the test. These will appear as extended periods of crazy/comically bad data. This requires a subjective judgment, see video for tips. Consult the lab notebook to see if notes were made about this channel during testing.

**2. Interpolate removed channels**

For any channels that have been removed, create a new “virtual” channel in its place that consists of the average activity from surrounding channels (to choose surrounding channels, consult a sensor distribution map of the net that was used; see video for details).

When naming interpolated channels, be sure to give them the same name as the channel you removed. For example, if you removed the channel ‘FCz’, rename the new interpolated channel ‘FCz’ as well. This way, if you apply a history template across all subjects, the new channel will still be recognized.

**3. Filter**:

Filtering removes very slow-moving waves (drifts) and very high frequency activity from the data. This “cleans up” the data and makes ERPs look smoother. For ERP analyses, apply a high-pass filter at 0.1 Hz, and a low-pass filter at 30 Hz.

For time-frequency analyses, apply a high-pass filter at 0.1 Hz, a notch filter at 60 Hz (to remove power-line interference), and a low-pass filter at 100 Hz.

**4. Reference**

Re-reference the data to the ‘average’ EEG signal. An average reference consists of all channels **except** the eye channels, which are **not** included.

Apply reference to all channels.

**5. Epoch (Segmentation)**

This step cuts the data up into time segments of EEG activity that are relevant to your task. Segment your data according to the reference markers (triggers) relevant to your stimuli. For this step, you may consider both the stimulus provided to the participant, as well as the response of the participant- each of which have their own marker (see video for details).

For ERP analyses use a -200ms pre-stimulus period, and a post-stimulus period long enough to capture any responses made by the participants (i.e., 1000ms).

For time frequency analysis, your epoch will be substantially larger, particularly in you are interested in slow-moving waves like delta. Remember that one cycle of delta takes 300-1000 ms. You need to epoch your data so that you have at least 3 cycles of the lowest frequency activity you plan on examining, or at least 1s before and after the time of interest (i.e., -1.5 – 3s if you want to examine -.5 to 2s).

**6. Ocular Correction**

Using either ICA or the ‘Gratton & Coles’ method, correct for any blinks made by the participants during the epochs.

If using ICA, (1) detect blinks using the mean sloped algorithm, (2) perform automatic mode selection, (3) use ‘whole data’ if computationally feasible (otherwise use interval suggested by BVA), (4) use the ‘Fast ICA’ algorithm, and (5) use 30% of relative variance as a similarity threshold. When asked, be sure to only compute as many components as you have independent recording sites (number of channels excluding interpolated channels).

**7. Baseline Correction**

This step normalizes activity within the time period you are interested in, relative to a time period when there is no task-relevant activity. For ERPs, use a period of -200ms to 0ms, baseline correct the data.

For time frequency-analysis, use a period of -500 to -300 ms.

**8. Artifact Rejection**

Using ‘Automatic Segment Selection’, apply artifact rejection to **all** channels.

Choice of rejection criteria will vary depending on the quality of the data (i.e., children, who typically have noisier data, might require less stringent criteria). Consider using both a Max-Min criterion and an Amplitude criterion.

As a start, for Max-Min, try allowing an absolute difference of 200 microvolts over 200ms. Mark 200ms both before and after the event as ‘bad’.

For Amplitude, try a minimal allowed amplitude of -100 microvolts, and a maximum allowed amplitude of 100 microvolts. Mark 200ms both before and after the event as ‘bad’.

Following these criteria, check to see how many segments were removed. Ideally, no more than a few segments (judgment required) should be removed from each condition. If a large number of segments have been rejected, consider reducing the stringency of either criteria.

**9. Average**

If you plan on ERP analyses, then average segments within each condition. This creates a “single-subject” average.

If you plan on time-frequency analyses, export data to MATLAB using a ‘Generic Data Export’. Change data from a ‘.dat’ file to a ‘.seg’ file. Include both a ‘Header’ and ‘Marker’ file. Data should be ‘Binary’, ‘Multiplexed’, and ‘IEEE 32-Bit Floating-Point Format”.