

ERP Compendium of Open Resources and Experiments (CORE)

MMN Analysis Procedures

Emily S. Kappenman^{1,2}, Jaclyn L. Farrens¹, Wendy Zhang^{1,2}, Andrew X. Stewart³, & Steven J. Luck³

¹ San Diego State University, Department of Psychology, San Diego, CA, 92120

² SDSU/UC San Diego Joint Doctoral Program in Clinical Psychology, San Diego, CA, 92120

³ University of California, Davis, Center for Mind & Brain and Department of Psychology, Davis, CA, 95616

This document outlines data processing procedures for isolating and measuring the mismatch negativity (MMN) from the passive auditory oddball paradigm of the ERP CORE (Compendium of Open Resources and Experiments) using the CORE data set and analysis files. The experiment control scripts, raw data, analysis scripts, processed data, and a broad set of results are available in the online resource (<https://doi.org/10.18115/D5JW4R>). Additional ERP resources are available on our website (<https://erpinfo.org>).

All analyses are performed in MATLAB using the EEGLAB (<https://sccn.ucsd.edu/eeglab/download.php>) and ERPLAB (<https://github.com/lucklab/erplab/releases>) open source toolboxes. The versions of EEGLAB (v.13.4.4b) and ERPLAB (v.8.0.0) that were used in conducting the original analyses are available for download with the CORE materials. Other versions of the toolboxes may work but have not been tested. Further details of the participants, stimuli, task, analyses, and results can be found in the accompanying ERP CORE publication (Kappenman et al., under review).

Rather than having one very long analysis script, we have divided the processing steps across several scripts. They should be run in the order listed below. In general, we recommend running whole scripts (e.g., by opening the script and clicking the Run icon) rather than selecting individual lines of code and running them (which can cause problems with paths).

The scripts and associated processing files are contained in subfolders located within the main MMN directory (each script path is specified below). Each subject's data and output files are contained in their individual subject folder located within the main MMN directory.

Individual-Subject EEG and ERP Processing Procedures

Script #1: Import_Raw_EEG_DS_Reref_Hpfil.m

This script (located in .../MMN/EEG_ERP_Processing) performs initial processing on the raw continuous EEG data file using the following operations:

1. Load the continuous raw EEG data file
 - The raw EEG data files already have been converted from their original file format to the .set EEGLAB file format. If you are applying this script to your own data, you will need to import your data into EEGLAB before running this script (see the EEGLAB documentation for more information on importing EEG files) and adjust the file names and paths to match your data.
2. Downsample the data from 1024 Hz to 256 Hz to speed data processing
 - This function automatically applies the appropriate low-pass anti-aliasing filter for the new sampling rate.
3. Re-reference the data to the average of P9 and P10 (located adjacent to the mastoids); create a bipolar HEOG channel (HEOG_left minus HEOG_right) and a bipolar VEOG channel (VEOG_lower minus FP2)
 - Bipolar EOG signals are particularly helpful in identifying ocular artifacts and will be used during artifact rejection (see Luck, 2014 for more details).
4. Add 3-D channel location information corresponding to the International 10-10 System
 - A channel location file that is appropriate for our recorded data set is included in the downloaded materials (standard-10-5-cap385.elc). Additional channel locations files can be found by consulting the help page for pop_chanedit.
5. Remove the DC offsets and apply a high-pass filter (non-causal Butterworth impulse response function, half-amplitude cut-off at 0.1 Hz, 12 dB/oct roll-off)
 - These data were acquired with an EEG system that records at DC (i.e., uses no high-pass filter during acquisition).

Script #2: ICA_Prep.m

This script (located in .../MMN/EEG_ERP_Processing) uses the output from Script #1 and prepares the EEG for computing ICA weights using the following operations:

1. Load the continuous EEG data file created by Script #1
2. Remove segments of EEG during the break periods between trial blocks (defined as 2 seconds or longer in between successive stimulus event codes)
 - Break periods tend to contain high levels of noise; therefore, we remove them prior to computing ICA weights.
3. Load parameters for rejecting especially noisy segments of EEG during trial blocks from the Excel file named ICA_Prep_Values_MMN.xlsx
 - Default parameters were used initially but were modified on an individual-subject basis where necessary based on visual inspection of the data.
4. Delete segments of the EEG exceeding the thresholds defined in the above Excel file

- This gets rid of especially noisy data that may interfere with the ICA process, but does not delete “ordinary” artifacts (e.g., eyeblinks and eye movements).

Script #3: Run_ICA.m

This script (located in .../MMN/EEG_ERP_Processing) uses the output from Script #2 and computes the ICA weights using the following operations:

IMPORTANT: The results of ICA decomposition (i.e., the ordering of the components, the scalp topographies, and the time courses of the components) will differ slightly each time the ICA weights are computed. This is because ICA decomposition starts with a random weight matrix (and randomly shuffles the data order in each training step), so the convergence is slightly different every time it is run. As a result, the topographic maps of the ICA weights and the Excel spreadsheet (ICA_Components_MMN.xlsx) containing the list of ICA component(s) to be removed for each subject included in this package will no longer be valid if you run the ICA decomposition. Consequently, to avoid confusion or accidental overwriting of relevant data files, this script has been commented out. You must skip the steps in this script and proceed directly to Script #4 if you want the results you get from these scripts to *exactly* match our results. Alternatively, this script can be un-commented and the ICA weights can be re-computed, but note that in that case the ICA component(s) will need to be re-chosen and the new values will need to be entered into the Excel file prior to running Script #4.

1. Load the EEG data file created by Script #2
2. Compute ICA weights with binICA (a compiled and faster version of ICA)
 - Note that the bipolar HEOG and VEOG channels are not included in the channel list for computing ICA weights, because they are not linearly independent of the channels that were used to create them. However, we will use these channels at a later stage so that we can see if blinks or eye movements were present that may have interfered with the perception of the stimuli.
 - If binICA is not an option (e.g., on a Windows machine), use runICA (see the script for more details).
3. Load the continuous EEG data file created by Script #1 and transfer ICA weights computed above to the continuous EEG data file (e.g., the version that did not have the break periods and noisy segments removed)
 - The ICA weights are transferred back to the continuous EEG data file created by Script #1 to ensure that breaks in the EEG resulting from the removal of sections of the EEG are not treated as continuous in the analysis.
4. Save a PDF of the topographic maps of the ICA weights for later review

Script #4: Remove_ICA_Components.m

This script (located in .../MMN/EEG_ERP_Processing) uses the output from Script #3 and removes the ICA component(s) corresponding to ocular artifacts (identified by visual inspection) that are listed in the Excel spreadsheet using the following operations:

1. Load the EEG data file containing the ICA weights created by Script #3

2. Load list of ICA component(s) corresponding to ocular artifacts from Excel file ICA_Components_MMN.xlsx
 - These component(s) were identified on the basis of visual inspection of the component time courses and topographic maps to correspond to ocular artifacts.
 - If you apply these scripts to your own data, you will need to determine which ICA components should be removed for each individual subject and put that information into the Excel file.
3. Perform ocular correction by removing the ICA component(s) specified in the Excel file above
4. Create a bipolar HEOG channel (HEOG_left minus HEOG_right) and a bipolar VEOG channel (VEOG_lower minus FP2) from the ICA-corrected data
 - These bipolar channels provide a way of seeing how well the artifact correction worked.
 - The original uncorrected bipolar HEOG and VEOG channels are also retained for subsequent artifact rejection procedures.
5. Add channel location information corresponding to the 3-D coordinates of the electrodes based on 10-10 International System site locations
 - Channel locations are re-calculated, because the channel operations procedure to create the ICA-corrected bipolar EOG channels removes the original channel location information. The Channel Operations tool always removes the electrode locations, because it has no way of knowing whether the locations are still valid after the operations have been performed.

Script #5: Elist_Bin_Epoch.m

This script (located in .../MMN/EEG_ERP_Processing) uses the output from Script #4 and epochs the data using the following operations:

1. Load the ICA-corrected EEG data file created by Script #4
2. Create EEG Event List containing a record of all event codes and their timing
 - The event list is stored in a text file (MMN_Eventlist.txt) and can be reviewed using a text editor.
3. Assign events to bins with Binlister
 - An individual trial may be assigned to more than one bin. Bin assignments can be reviewed in each subject's MMN_Eventlist_Bins.txt file.
4. Epoch the EEG into 1-second segments time-locked to the stimulus (from -200 ms to 800 ms), and perform baseline correction using the average voltage from -200 ms to 0 ms

Script #6: Artifact_Rejection.m

This script (located in .../MMN/EEG_ERP_Processing) uses the output from Script #5 and performs artifact rejection using the following operations:

1. Load the Excel files with parameters for artifact rejection tailored to each individual subject's data (see details below)
 - Default parameters were used initially and modified as necessary for each individual subject.

- If you apply these scripts to your own data, you will need to visually inspect each subject's data and modify these files to contain appropriate values.
2. Interpolate channel(s) specified in Excel file Interpolate_Channels_MMN.xlsx
 - Channel(s) to interpolate were chosen by visual inspection of the data.
 - Any channels without channel locations (e.g., the EOG channels) should not be included in the interpolation process and are listed in ignored channels.
 - EEG channels that will later be used for measurement of the ERPs should not be interpolated; if a subject has an especially noisy measurement channel, that subject should be removed from the analysis. Consequently, the interpolation only influences scalp maps, not the primary amplitude and latency values that would be used in statistical analyses.
 3. Identify segments of EEG with C.R.A.P. using the simple voltage threshold algorithm with the parameters in the Excel file AR_Parameters_for_SVT_CRAP_MMN.xlsx
 - This operation is performed on all channels except for the bipolar HEOG and VEOG channels
 4. Identify segments of EEG with C.R.A.P. using the moving window peak-to-peak algorithm with the parameters in the Excel file AR_Parameters_for_MW_CRAP_MMN.xlsx
 - This operation is performed on all scalp EEG channels
 5. Identify segments of EEG with any uncorrected horizontal eye movements using the step like algorithm with the parameters in the Excel file AR_Parameters_for_SL_HEOG_MMN.xlsx
 - This operation is performed on the ICA-corrected HEOG channel
 - Although ICA component(s) are removed to correct for horizontal eye movements, ICA correction tends not to be as effective for horizontal eye movements as for eyeblinks. We therefore perform an artifact rejection procedure to identify and remove any epochs containing uncorrected horizontal eye movements.

Script #7: Average_ERPs.m

This script (located in .../MMN/EEG_ERP_Processing) uses the output from Script #6 and creates averaged ERP waveforms using the following operations:

1. Load the epoched and artifact rejected EEG data file created by Script #6
2. Create an averaged ERP waveform
 - The epochs marked as containing artifacts in Script #6 are excluded from the average.
 - A separate averaged ERP waveform is created for each bin.
3. Apply a low-pass filter (non-causal Butterworth impulse response function, 20 Hz half-amplitude cut-off, 48 dB/oct roll-off) to the ERP waveforms
 - The low-pass filtered ERP waveforms will be used subsequently for plotting and for calculation of ERP measurements that are susceptible to influence from high frequency noise (i.e., peak amplitude, peak latency, and onset latency).
4. Calculate the percentage of trials rejected (per bin and in total) and save the information to a .csv file in each subject's data folder
 - This information should be reviewed for each subject to ensure that the percentage of trials rejected is within the accepted range. In our studies of neurotypical young adults, we always exclude any subject with 25% or more trials rejected.

5. Calculate ERP difference waveforms between conditions and create low-pass filtered versions of the ERP waveforms.

Script #8: Plot_Individual_Subject_ERPs.m

This script (located in .../MMN/EEG_ERP_Processing) uses the output from Script #7, plots the relevant ERP waveforms, and saves PDFs of the plots using the following operations:

1. Set parameters for plotting the individual subject ERP waveforms, including baseline correction period and scales for the x-axis (time, in milliseconds) and y-axis (amplitude, in microvolts)
 - Setting a uniform scale for plotting individual subject ERPs makes it easy to compare ERP waveforms across subjects.
2. Load the low-pass filtered averaged ERP waveform created by Script #7
3. Plot the MMN deviant and standard parent waveforms at the electrode sites of interest and save a PDF of the plot in the subject's folder
4. Plot the MMN deviant and standard parent waveforms at all electrode sites and save a PDF of the plot in the subject's folder
5. Plot the MMN deviant-minus-standard difference waveform at the electrode sites of interest and save a PDF of the plot in the subject's folder
6. Plot the MMN deviant-minus-standard difference waveform at all electrode sites and save a PDF of the plot in the subject's folder
7. Plot the ICA-corrected and uncorrected bipolar HEOG signals, and save a PDF of the plot in the subject's folder
 - The HEOG plots should be reviewed to ensure that no residual artifacts remain.
8. Plot the parent ICA-corrected monopolar VEOG signals and corrected bipolar VEOG signal, and save a PDF of the plot in the subject's folder
 - The VEOG plots should be reviewed to ensure that no residual artifacts remain in the ICA-corrected waveforms, with particular focus on potential polarity inversions below versus above the eye.

Grand Average ERP Processing Procedures

Script #9: Grand_Average_ERPs.m

This script (located in .../MMN/EEG_ERP_Processing/Grand_Average_ERPs) uses the output from Script #7 and creates grand average ERP waveforms using the following operations:

1. Specify the list of subjects to include in the grand average ERP waveforms (i.e., subjects who were not excluded due to excessive artifacts)
2. Create grand average ERP waveforms from individual subject ERPs without low-pass filter applied
3. Create grand average ERP waveforms from individual subject ERPs with a low-pass filter applied

Script #10: Plot_Grand_Average_ERPs.m

This script (located in .../MMN/EEG_ERP_Processing/Grand_Average_ERPs) uses the output from Script #9, plots the grand average ERP waveforms, and saves PDFs of the plots using the following operations:

1. Set parameters for plotting the individual subject ERP waveforms, including baseline correction period and scales for the x-axis (time, in milliseconds) and y-axis (amplitude, in microvolts)
2. Load the low-pass filtered averaged ERP waveform created by Script #9
3. Plot the MMN deviant and standard parent waveforms at the electrode sites of interest and save a PDF of the plot in the Grand_Average_ERPs folder
4. Plot the MMN deviant and standard parent waveforms at all electrode sites and save a PDF of the plot in the Grand_Average_ERPs folder
5. Plot the MMN deviant-minus-standard difference waveform at the electrode sites of interest and save a PDF of the plot in the Grand_Average_ERPs folder
6. Plot the MMN deviant-minus-standard difference waveform at the electrode sites of interest with the standard error of the mean (SEM) and save a PDF of the plot in the Grand_Average_ERPs folder
7. Plot the MMN deviant-minus-standard difference waveform at all electrode sites and save a PDF of the plot in the Grand_Average_ERPs folder.
8. Plot the parent (deviant and standard trials) ICA-corrected and uncorrected bipolar HEOG signals, and save a PDF of the plot in the Grand_Average_ERPs folder
9. Plot the parent (deviant and standard trials) ICA-corrected monopolar VEOG signals and corrected bipolar VEOG signal, and save a PDF of the plot in the Grand_Average_ERPs folder

Script #11: Plot_Grand_Average_Topomaps.m

This script (located in .../MMN/EEG_ERP_Processing/Grand_Average_ERPs) uses the output from Script #9, plots topographic maps of the mean amplitude of the components, and saves PDFs of the plots using the following operations:

1. Set parameters for plotting the topographic maps of the grand average ERPs, including

- Mean amplitude time window in milliseconds (e.g., 125 to 225 ms)
 - EEG channels to include in the topographic maps
 - Bins to create topographic maps for (a separate topographic map will be created for each bin specified)
 - Optional color scale limits for the topographic maps for each bin (in microvolts)
2. Load the grand average ERP waveforms (without low-pass filter applied) in .erp ERPLAB file format created by Script #9
 3. Calculate the mean amplitude for each bin and channel based on the specified time window
 4. Create a topographic map of the mean amplitude for each specified bin, and save a PDF of the plots in the Grand_Average_ERPs folder

ERP Measurement Procedures

Script #12: Measure_ERPs.m

This script (located in .../MMN/EEG_ERP_Processing/ERP_Measurements) uses the output from Script #7 and measures the amplitudes and latencies of the ERPs using the following operations:

1. Specify the list of subjects to include when measuring the ERP waveforms (i.e., subjects who were not excluded due to excessive artifacts)
2. Set parameters for measuring ERPs, including:
 - Measurement time window for measuring mean amplitude, peak amplitude, peak latency, and 50% area latency in milliseconds (e.g., 125 to 225 ms)
 - Measurement time window for measuring onset latency (50% peak latency) in milliseconds (e.g., 25 to 225 ms)
 - EEG channel(s) to measure the components
 - Difference wave and parent wave bins for measurement
 - Baseline correction period for measurement in milliseconds
3. Calculate difference waveform measurements on averaged ERP waveforms without a low-pass filter applied
 - Create a text file containing a list of unfiltered ERPsets and their file locations
 - Measure mean amplitude using the time window, channel(s), and difference wave bin(s) specified in Step 2
 - Measure 50% area latency (negative area only) using the time window, channel(s), and difference wave bin(s) specified in Step 2
 - Note: low-pass filtering is unnecessary for mean amplitude and 50% area latency because these measures are not distorted much by high-frequency noise
4. Calculate difference waveform measurements on averaged ERP waveforms with a low-pass filter applied
 - Create a text file containing a list of low-pass filtered ERPsets and their file locations
 - Measure local peak amplitude using the time window, channel(s), and difference wave bin(s) specified in Step 2
 - Measure local peak latency using the time window, channel(s), and difference wave bin(s) specified in Step 2
 - Measure onset latency (50% peak latency) using the time window, channel(s), and difference wave bin(s) specified in Step 2
5. Parent waveform measurements on averaged ERP waveforms without a low-pass filter applied
 - Measure mean amplitude using the time window, channel(s), and parent wave bin(s) specified in Step 2
6. For each measurement described above, a text file containing measurement values for each subject is saved in the ERP Measurements folder

Script #13: Plot_Measurement_Histograms.m

This script (located in .../MMN/EEG_ERP_Processing/ERP_Measurements) uses the output from Script #12 and creates histograms of the individual subject measurement values using the following operations:

1. Specify the measurements that will be used to create histograms using the text files created by Script #12 (e.g., Mean_Amplitude_Diff_Waves.txt)
2. For each measurement listed in Step 1, create a histogram plot of the single-participant measurement values, and save PDFs of the histograms in the ERP_Measurements folder

EEG and ERP Noise Measurement Procedures

Script #14: FFTs.m

This script (located in .../MMN/Noise_Measurements) uses the output from Script #2 and computes the fast Fourier transform (FFT) using the following operations:

1. Load the EEG data file with break periods and excessively noisy segments removed created by Script #2
2. Compute FFT on the EEG channel(s) of interest for each subject averaged across 5-second moving-window segments of the EEG (with 50% overlap)
3. Average FFTs across all subjects and plot the grand average FFT. A PDF of the plot is saved in the Noise_Measurements folder.

Script #15: Baseline_Noise.m

This script (located in .../MMN/Noise_Measurements) uses the output from Script #7 and calculates the noise in the baseline period of the ERP waveform using the following operations:

1. Set parameters for calculating noise in the baseline period for each subject, including the baseline period (in milliseconds) and channel(s) and bin(s) to measure baseline noise from
2. For each subject, load the averaged ERP waveforms with no low-pass filter applied created by Script #7
3. Calculate the baseline noise, defined as the standard deviation of the voltages in the baseline period, using the time window, channel(s), and bin(s) specified in Step 1 and write all individual subject's baseline noise values to a .csv file
4. Repeat steps 2-3 on the averaged ERP waveforms with a low-pass filter applied outputted from Script #7
5. Create baseline window noise probability histograms for the unfiltered and filtered plus-minus averages separately, overlaying the parent and difference conditions, and save PDFs of the plots in the Noise_Measurements folder

Script #16: PlusMinus_Averages.m

This script (located in .../MMN/Noise_Measurements) uses the output from Script #6, creates plus-minus average waveforms, and plots the plus-minus waveforms using the following operations:

1. For each subject, load the epoched and artifact rejected EEG data file created by Script #6
2. Find the odd-numbered and even-numbered epochs not marked for artifact rejection for each parent bin of interest (e.g., deviant trials and standard trials; the difference bin will be re-calculated below)
3. Create separate averaged ERP waveforms for the odd-numbered and even-numbered trials identified above
4. Create difference waveforms between conditions (e.g., deviant-minus-standard) separately for odd-numbered and even-numbered waveforms

5. Create plus-minus average waveforms by subtracting the ERP for even-numbered trials from the ERP for odd-numbered trials (i.e., odd minus even) and divide by 2
 - This is equivalent to inverting the polarity of the even-numbered trials and averaging them with the odd-numbered trials.
6. Apply a low-pass filter to the plus-minus waveforms (non-causal Butterworth impulse response function, 20 Hz half-amplitude cut-off, 48 dB/oct roll-off)
7. Append all of the individual subject low-pass filtered plus-minus average waveforms
8. Plot individual subject plus-minus waveforms overlaid on a single plot, separately for each condition. A PDF of the plot is saved in the Noise_Measurements folder.
9. Calculate and plot the standard deviation across subjects at each time point in the plus-minus average waveforms. A PDF of the plot is saved in the Noise_Measurements folder.

Script #17: Measurement_Window_Noise.m

This script (located in .../MMN/Noise_Measurements) uses the output from Script #16 and calculates the noise in the measurement window of the plus-minus averages using the following operations:

1. Set parameters for calculating noise in the measurement time window of the ERP component for each subject, including the measurement time window (in milliseconds) and channel(s) and bin(s) to measure noise from
2. For each subject, load the plus-minus average waveforms with no low-pass filter applied created by Script #16
3. Calculate the noise during the measurement time window, defined as the standard deviation of the voltage in the measurement window, using the time window, channel(s), and bin(s) specified in Step 1, and write values to a .csv file in the Noise_Measurements folder
4. Repeat steps 2-3 on the plus-minus averages with a low-pass filter applied
5. Create measurement window noise probability histograms for the unfiltered and filtered plus-minus averages separately, overlaying the parent and difference conditions, and save PDFs of the plots in the Noise_Measurements folder