Steve's Tips for Reading and Evaluating ERP Studies

For more details, see Chapter 15 in Luck (2014): Reading, Writing, and Reviewing ERP Papers (available [online](#)).

Before you look at the data…

Before you look at the data in an ERP study, you should first check out the filter settings and the reference electrode.

Unless you know what the researchers used as the reference, you don't know what you're really seeing in the data from a given channel. For example, the channel labeled "Pz" might look completely different depending on whether the reference was Cz, the average of the mastoids, or the average of all electrode sites.

And if the researchers used the average of all electrodes as the reference, they really ought to explicitly list all of the sites. Sometimes a paper will just say something like "We recorded data from 32 standard 10/20 electrodes, including the PO3 and PO4 sites that were the focus of our analyses." If they used the average of all sites as the reference, but they didn't tell you what sites they included in their recordings, you don't really know what you're looking at.

The table to the right shows my recommendations for evaluating filter settings in most research on cognitive and affective research with adults. Different settings might be appropriate for early sensory components or for other participant populations.

Sometimes the settings labeled "worry a lot" are actually okay in a given study, but the authors should probably justify these settings. Also, the filters may have significantly impacted the onset and offset times of the effects (for a low-pass cutoff of < 10 Hz) and may have created spurious peaks in the waveforms (for a high-pass cutoff of > 0.5 Hz).

|	High-pass cutoff	Low-pass cutoff
Don't worry	≤ 0.1 Hz	≥ 20 Hz
Worry a little	0.1-0.5 Hz	10-20 Hz
Worry a lot*	> 0.5 Hz	< 10 Hz

*Especially when slope is > 12 dB/octave

When looking at ERP waveforms…

Before you look at the waveforms, you absolutely must determine the location used as the reference (see the previous section).

The first thing you should do when looking at ERP waveforms is inspect the baseline period. Do they not show a baseline period? That's a warning sign.

When you look at a prestimulus baseline period, the goal is to evaluate the noise level and look for evidence of differential overlap/preparation between the conditions. If the noise is in the baseline is not substantially smaller than the effects described in the paper, you should be suspicious that the effects are also noise (even if they're statistically significant).

If you see different "tilts" in the baseline for the conditions or groups being compared (as in the waveforms shown here), you should be concerned about differential overlap/preparation. In addition, if you see effects that begin unrealistically early (e.g., < 100 ms for most cognitive effects) and last for a long time, that is also a common signature of baseline problems (e.g., noise, differential overlap, differential preparation).
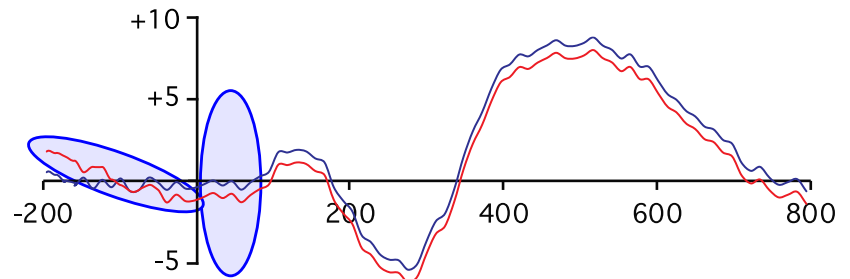
Top 10 problems

The following are the 10 most common problems I see when I read and review ERP papers. You should look for them in every ERP paper you read.

Data problems

1) Noisy data- Don't believe an effect that is smaller than the baseline noise, even if it's statistically significant (unless it has been replicated).

2) Baseline problems- Noise, differential overlap, or preparatory effects are often visible in the baseline period.

3) Blinks or eye movements- An effect that is largest at the frontal poles might be a blink artifact. This can be ruled out if the effect does not invert under the eyes. Eye movements in the direction of the target or the response hand are a major problem in studies looking at lateralized components (e.g., N2pc, CDA, LRP). These studies should carefully demonstrate that the results are not contaminated by lateral eye movements.

Analysis problems

4) Inappropriate filtering- See the table on the previous page.

5) Inappropriate amplitude or latency measures- Peak measures can be problematic, especially if the waveforms are noisier in one group/condition than in the others (e.g., due to different numbers of trials per average). However, there is usually no need to equate the number of trials when mean amplitude is measured instead of peak amplitude.

6) Statistical problems- It can be a major problem if the researchers look at the data and use the observed effects to decide on the time windows and electrode sites to use in their analyses. This leads to a major inflation of the Type I error rate (the rate of false positives). If the measurement parameters are not chosen a priori, the researchers should explain how they avoided biasing the results when they chose the time windows and electrode sites for analysis.

Design and interpretation problems

7) Lack of specific predictions- The Introduction to the paper should provide specific predictions about the ERPs that would be expected according to the hypotheses being tested.

8) Physical stimulus confounds- If the stimuli differ across the groups or conditions being studied, this may explain the observed differences in ERPs (especially for effects within 200-300 ms of stimulus onset). Don't forget the Hillyard Principle: "To avoid physical stimulus confounds, use identical stimuli across conditions and vary only the task instructions."

9) Failure to isolate the component of interest- Did the researchers actually isolate the component of interest, or might the effects reflect a completely different neurocognitive process?

10) Overreliance on source localization- It can be useful for a paper to include information about the plausible neural generator sources of their effects as long as they are careful to say that the data are *consistent* with a particular generator source rather than that the data *demonstrate* that a particular part of the brain is involved. If the researchers want to draw strong conclusions about the generator source, they would need to provide a principled quantification of the accuracy of the solution.