I implemented various shuffle options on the data, to check both for unexpected structure in the data and possible bugs in our analysis. There were some surprises, which although not in the end indicative of any real problems in the analysis, are interesting enough that I thought I should write out what I found. This is mostly for deep background.

The shuffling may now be done with any combination of shuffling over trials and shuffling paint/shadow. The shuffling over trials is done first, separately for paint and shadow. One can shuffle trials within intensity (disrupts correlations across response channels) or over all trials (disrupts relation between stimulus intensity and response).

The paint/shadow shuffling preserves the intensity response relations, which themselves were either shuffled as described above or not.

The fullest shuffle is thus over all trials and then additionally over paint and shadow. We expect no structure in the decoding for this case. Shuffle Figure 1 shows the decoded ranges for all conditions for this case, for two independent runs of the shuffled analysis. The x-axis is the nominal session/condition number. Given that the range will always be positive, small non-zero values are expected. For comparison, we have been using 0.2 as our threshold for using a decoding in the main analysis. None of the shuffled decoded ranges exceed this threshold. As shown in Shuffle Figure 2, the decoded intensity plots from example single sessions are flat. So this seems about right to me.

A minor mystery, which I have not yet solved, is why the pattern of decoded ranges over sessions is always about the same. (Shuffle Figure 1 shows two independent runs, but I have examined this pattern over more runs and the plots always look about the same.) I conjecture that in some sessions there is more imbalance in the number of usable trials at each probe intensity, and that this causes the decoding regression to fit those intensities better than some other ones. I can imagine that this might produce systematically larger decoded ranges for some sessions than others. I am not sure this idea really holds water. I am also not sure whether this is worth tracking down (but it does bug me a little). The decoded range is computed as max decoded intensity minus mean decoded intensity, separately for paint and shadow, and then averaged for paint and shadow. It is on the leave-one-out predictions if a leave-one-out analysis was done (which it was here).

A bigger mystery, which I have now sorted out, is that it appears that the decoded intercepts can be systematically different from (less than) zero for the fully shuffled case. I first noticed this by looking at the summary decoded intercept plots and thinking that there were too many cases when the decoded intercepts were negative. So I set things up so I could re-run the shuffled analysis 10 times and did a t-test on the intercepts.

Decode intercept V4: -0.01 +/- 0.016, t-test different from 0: p = 0.600

Decode intercept V1: -0.11 +/- 0.024, t-test different from 0: p = 0.003

The intercept is significantly negative for the V1 data in this set, although not for the V4 data. This bugged me quite a bit -- how could there be a systematic effect.

It turns out that we introduce an asymmetry when we fit the intercept on specified intensities for the paint condition and infer the matched intensities for the shadow condition. Because the decoding is so bad, the inference of matches is very unstable. For many conditions, there is no overlapping decoded range and we simply can't infer an intercept. But in some cases (left panel of Shuffle Figure 3), the decoded curve runs horizontally (Shuffle Figure 3) and sometimes vertically (Shuffle Figure 4), based on very very small differences in the offset between the very very flat decoded intensities. (See faint green curve in left hand plots, which represents the full range of inferred matches.) In the case like Shuffle Figure 3, we get a negative decoded intercept. In the case like Shuffle Figure 4, there is no decoded intercept because no paint intensities in the range of the psychophysics (0.25-0.75) have an inferred match. So, basically, we tend to include cases with a negative intercept in the overall average than we do cases with a positive intercept, simply based on our selection of paint intensities. This pathology goes away if we instead get the average intercept of the line fit through all possible inferred matches.

Decode intercept smooth V4: 0.02 +/- 0.019, t-test different from 0: p = 0.399

Decode intercept smooth V1: -0.06 +/- 0.036, t-test different from 0: p = 0.127

We could also just look at the difference in average decoded intensity between paint and shadow (average height of the smooth lines in the left plots of Shuffle Figures 3 and 4). These are well defined for each condition and average out to zero.

Paint shadow decode diff V4: -0.00 +/- 0.000, t-test different from 0: p = 0.048

Paint shadow decode diff V1: -0.00 +/- 0.002, t-test different from 0: p = 0.950

So I think the shuffling is all working OK, and we are not extracting any surprising structure beyond that due to the asymmetry in intensity points chosen for the analysis.

Does this asymmetry cause us problems for our real analysis? I don't think so. It is a pathology caused by the flatness of the decoded intensities. To see this, note that if we just shuffle on paint/versus shadow, the intensity decoding remains good (effect not shown here) but the negative intercept bias goes away.

Decode intercept V4: -0.00 +/- 0.002, t-test different from 0: p = 0.339

Decode intercept V1: 0.00 +/- 0.007, t-test different from 0: p = 0.500

Decode intercept smooth V4: 0.00 +/- 0.002, t-test different from 0: p = 0.726

Decode intercept smooth V1: 0.01 +/- 0.006, t-test different from 0: p = 0.076

Paint shadow decode diff V4: 0.00 +/- 0.000, t-test different from 0: p = 0.343

Paint shadow decode diff V1: 0.00 +/- 0.001, t-test different from 0: p = 0.267

Also note (Shuffle Figure 5) that if we compare the selected intercept decoding and the intercept decoding based on the full set of decoded intensities for our real (unshuffled) decoding, they look essentially identical.

More interestingly than the above we can look at what happens if we shuffle trials within stimulus intensity, but leave paint/shadow labeling alone. Here we see that the decoded ranges are largely unaffected (Shuffle Figure 6) and the overall decodings also look about the same (Shuffle Figure 7). This suggests that across electrode/unit correlations are not carrying much information that we are using in the decoder.