Supplemental Materials

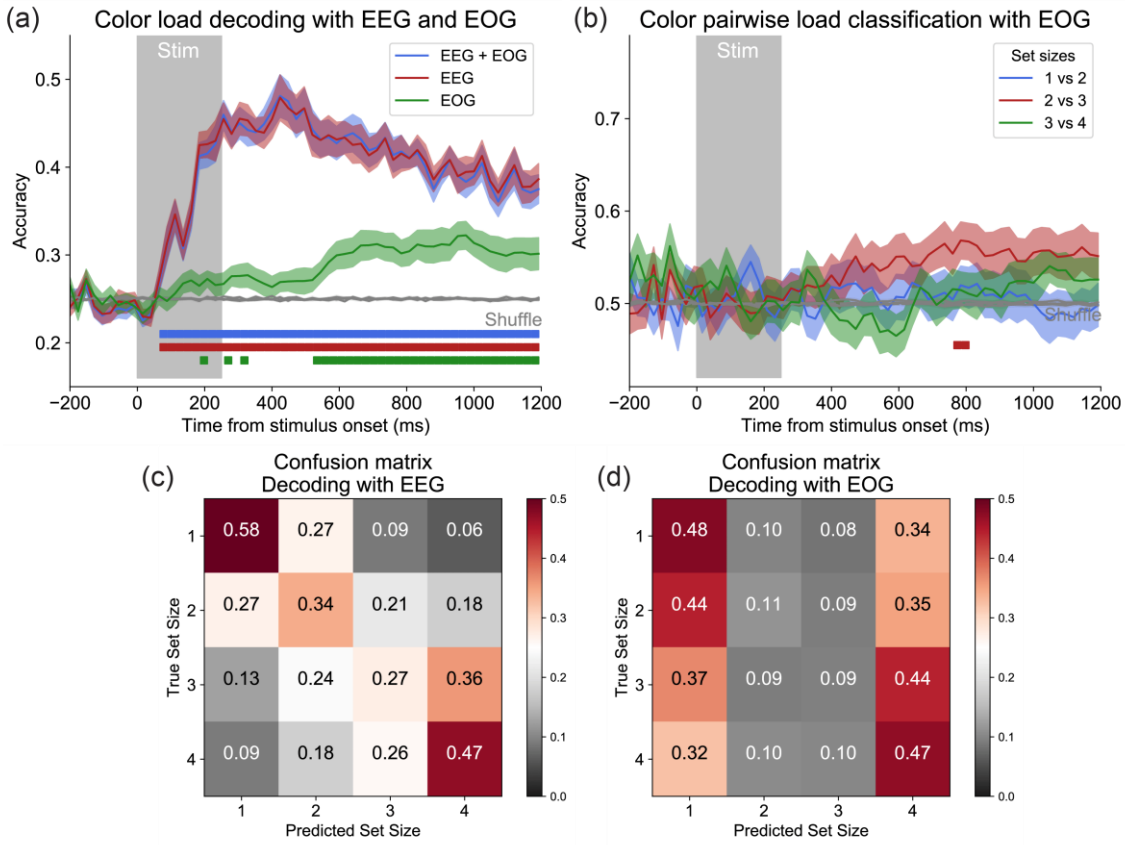

Figure S1 Load decoding with EOG data compared to EEG data. (a) Load decoding with EOG (in green) is numerically lower than with EEG (in red) while also consistently rising above chance much later in the delay period. Additionally, decoding load with both EOG and EEG (in blue) provides very similar decoding accuracy. (b) Decoding pairs of set sizes with EOG reveals very little sensitivity in detecting single-item additions to load. 1 versus 2 and 3 versus 4 are never significantly above chance. 2 versus 3 is only above chance for 2 timepoints. (c) The confusion matrix for load decoding with EEG shows a fairly strong diagonal, indicating that the set sizes are discriminable (errors are more likely to be towards adjacent set sizes). (d) The confusion matrix for load decoding with EOG shows no diagonal. Predictions are driven almost entirely by predicting the edge classes (1 and 4).

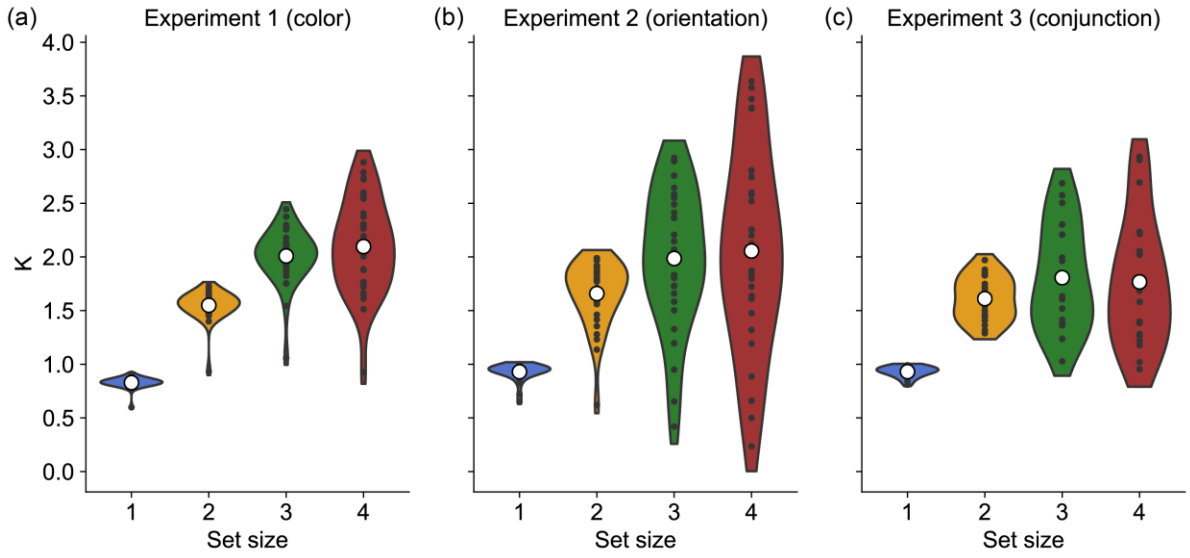

Figure S2 Working memory capacity (Cowan's k) for each set size in each of the three experiments.

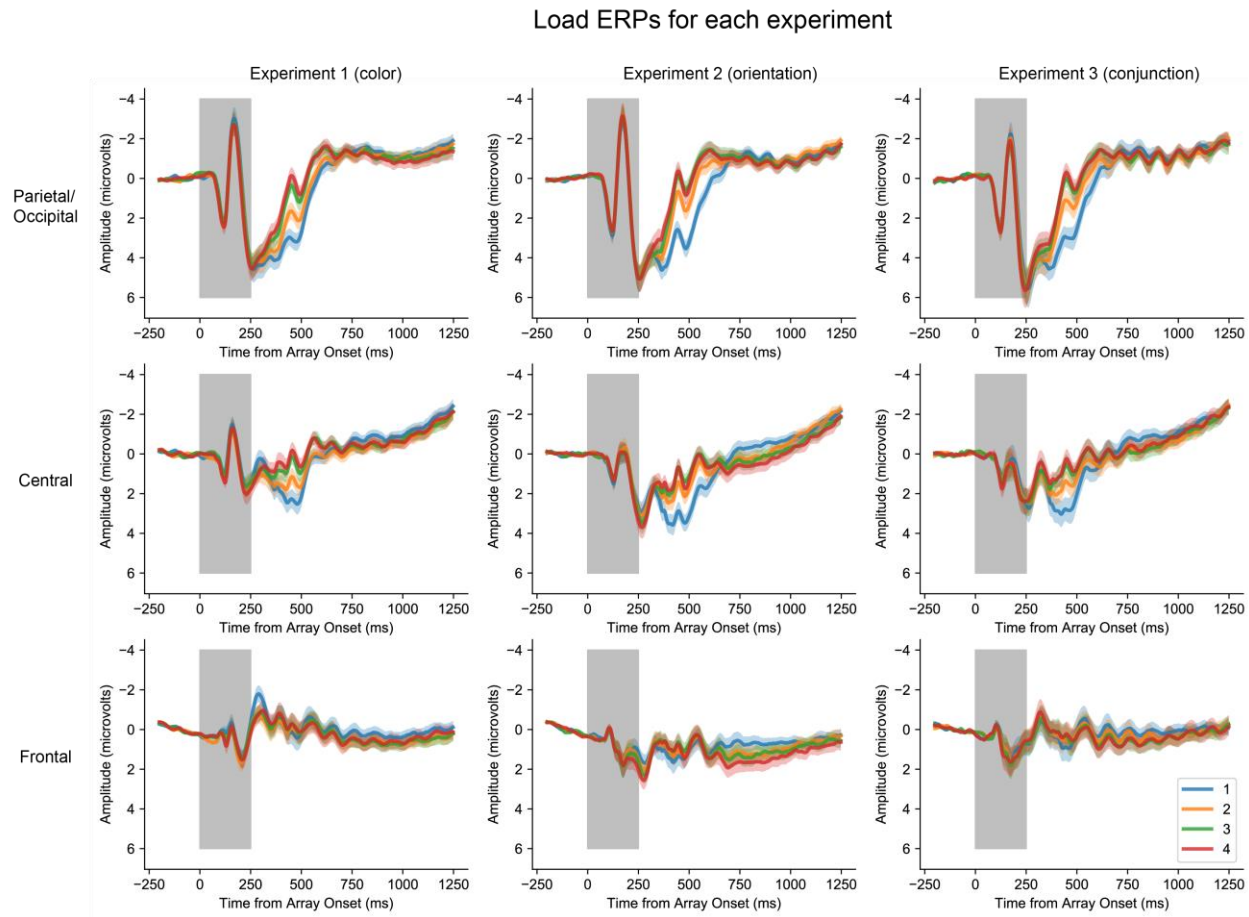

Figure S3 Load event related potentials (ERPs) for each experiment at parietal/occipital, central, and frontal electrodes.

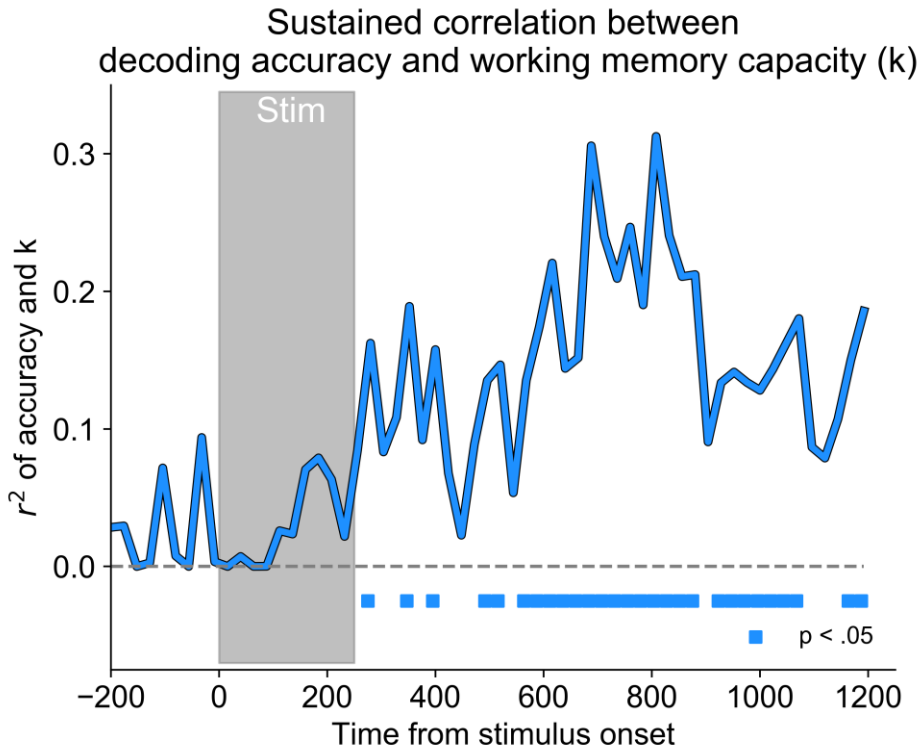

Figure S4 Time course of the coefficient of determination (r^2) between decoding accuracy and working memory capacity (Cowan's k). Blue squares indicate timepoints with significant r^2 value (corrected $p < .05$, *FDR* = .05 with Benjamini-Hochberg procedure).

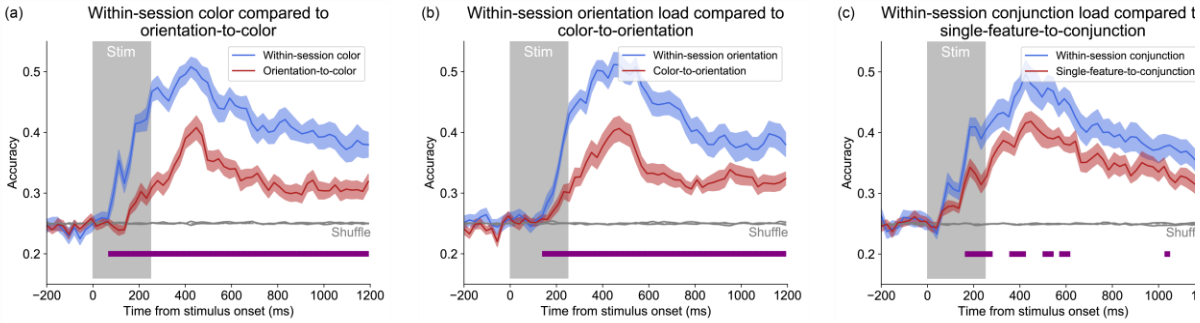

Figure S5 Comparing within-session (or within-experiment) load decoding to cross-training decoding. Purple squares indicate timepoints with within-session accuracy significantly above cross-training accuracy (corrected $p < .05$, *FDR* = .05 with Benjamini-Hochberg procedure). Accuracy in single-feature cross-training (a and b) is significantly lower than within-session accuracy for the entirety of the delay period. Accuracy single-feature to conjunction cross-training (c) is significantly lower than within-session accuracy, but only at 12 timepoints throughout the trial.

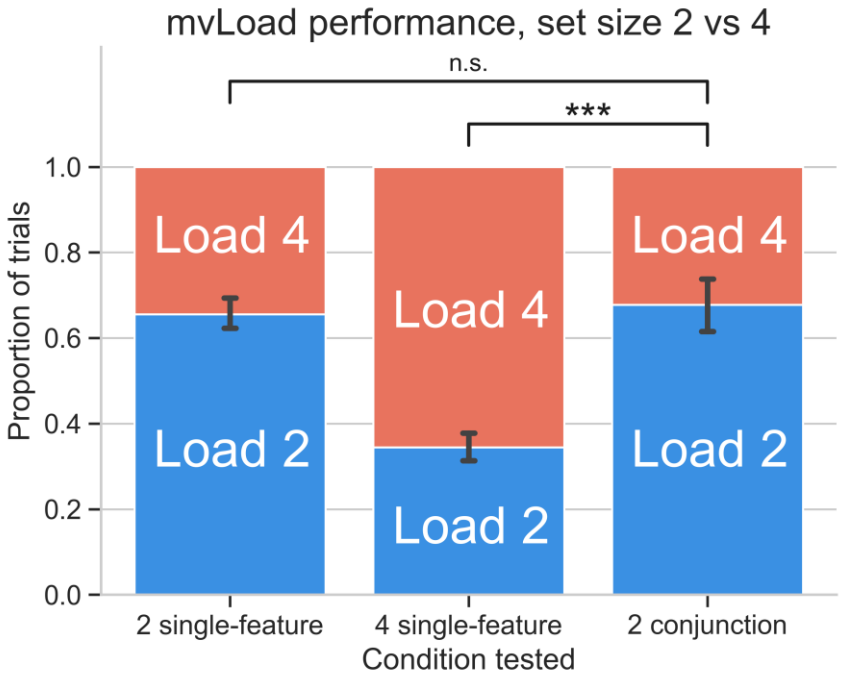

Figure S6 mvLoad classifier output (trained on single-feature items and tested on conjunction items). Each bar represents 100 percent of trials classified as load 2 (blue) or 4 (red) in each condition. Set size 2 conjunctions trials are predicted as set size 2 at the same rates as set size 2 single-feature trials.