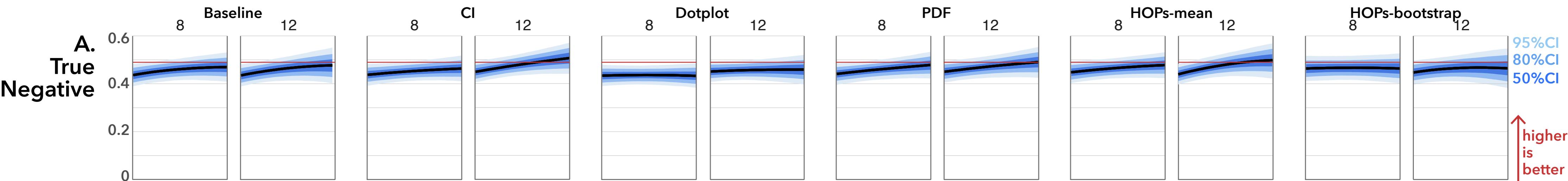
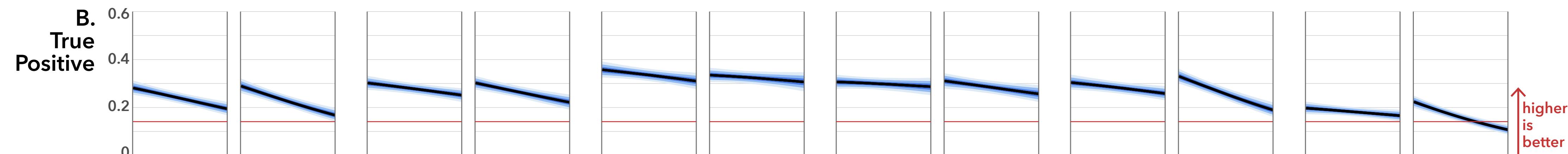


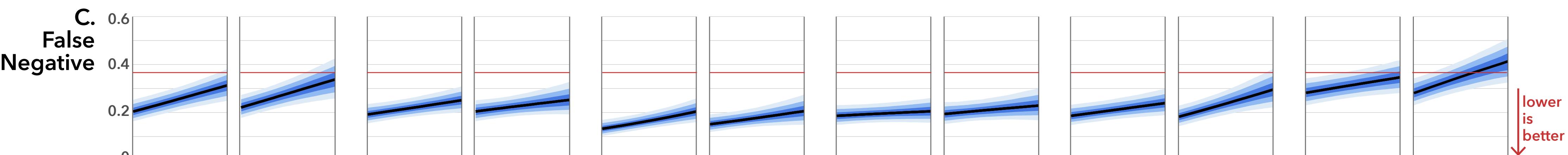
Our model predicts the proportion of True Negatives (TNs), True Positives (TPs), False Positives (FPs) and False Negatives (FNs) using (*uncertainty*) *display*, *number of graphs shown (nregion)* and *trial number* as predictors. Below we plot the average proportion, and 95% posterior credible intervals of TNs, TPs, FN and FPs for each *display* and *nregion* in a given trial, and its change as a typical participant progresses through the set of 70 trials (in two blocks).



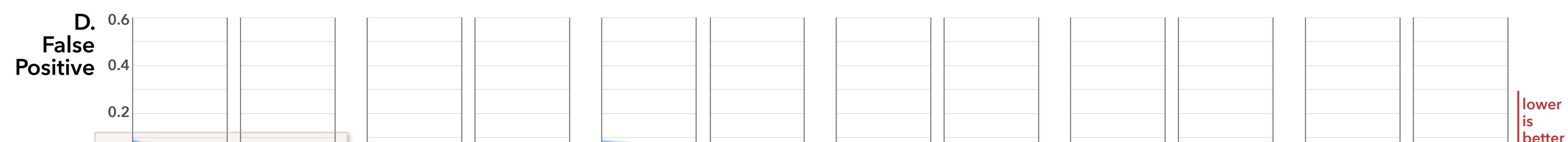
Proportion of FN increases as trials progress. This perhaps due to the strong incentives against FPs, resulting in participants being more cautious in selecting what they think are *positives*. The proportion of FN is higher when *nregion* = 12 compared to *nregion* = 8



Proportion of FPs decrease sharply over the course of trials, and continues to decrease after the initial five trials in each block where feedback is provided to them. Proportion of FPs are lower when *nregion* = 12 compared to *nregion* = 8

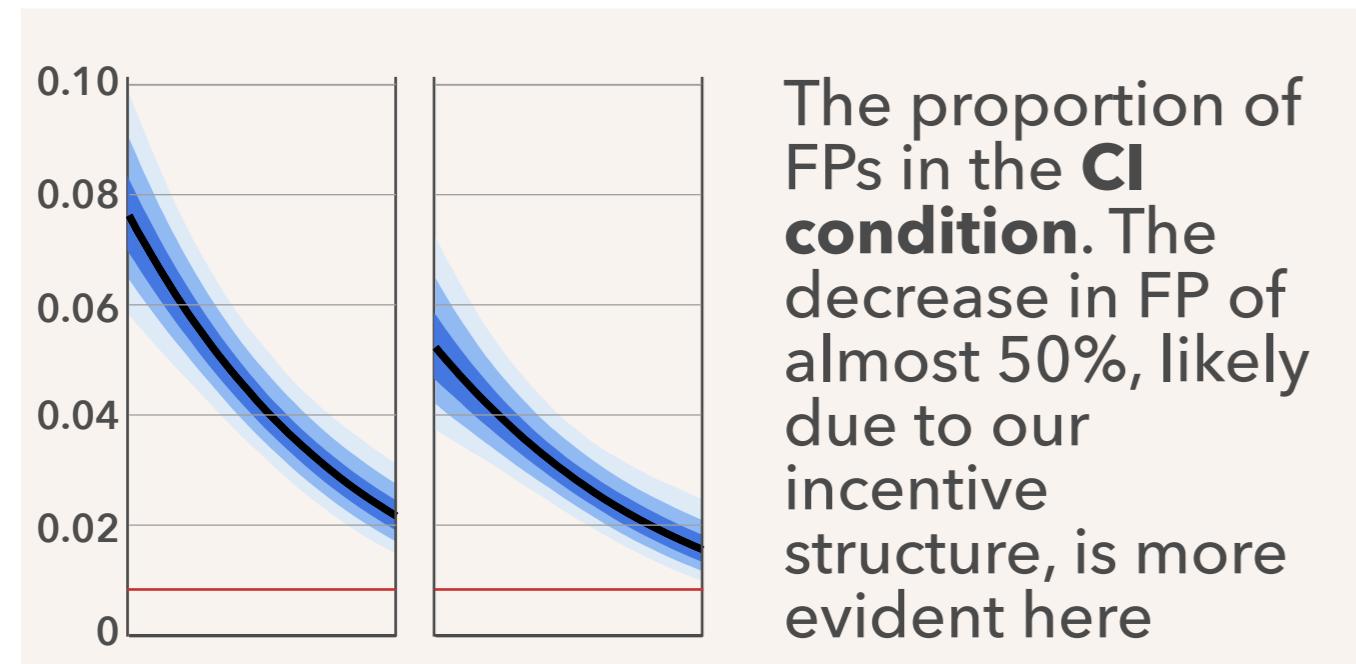


Proportion of TNs increase slightly over the course of trials in most conditions, although they are relatively constant for dotplots and HOPs-b. Proportion of TNs are usually higher when *nregion* = 12 compared to *nregion* = 8



Proportion of TPs decrease over the course of trials, and in some conditions, this decrease is quite sharp. In some conditions, the proportion of TPs is lower when *nregion* = 12 compared to *nregion* = 8

E. Closer look at proportion of False Positives



F. Calculation of marginalised density estimates (used in Figures 4 - 9)

