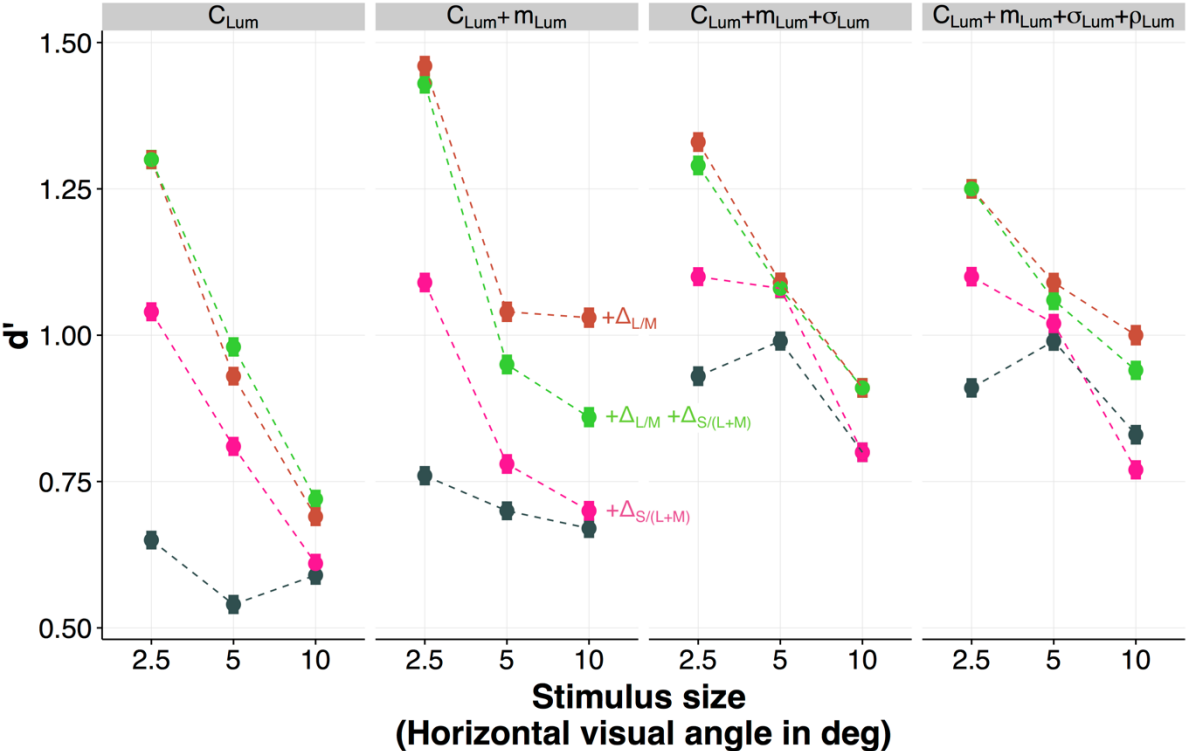


### S3 Figure: Classifier performances

We evaluated performance of our different classifiers for each image size independently. The  $d'$  values we computed are shown in Figure S3. Comparing these results with observers' performance in our psychophysical experiment, we see that almost all our Bayesian model classifiers fail in reproducing the effect of size observed for our participants.



**Figure S3** –Summary of classifiers performance expressed in terms of  $d'$  values. Intervals correspond to lower and upper 95<sup>th</sup> percentile confidence interval based on parametric bootstrap simulations (n=1000).