Accuracies on the M-out-of-N test set

