Accuracy Scores Human RTs vs. Model Surprisals