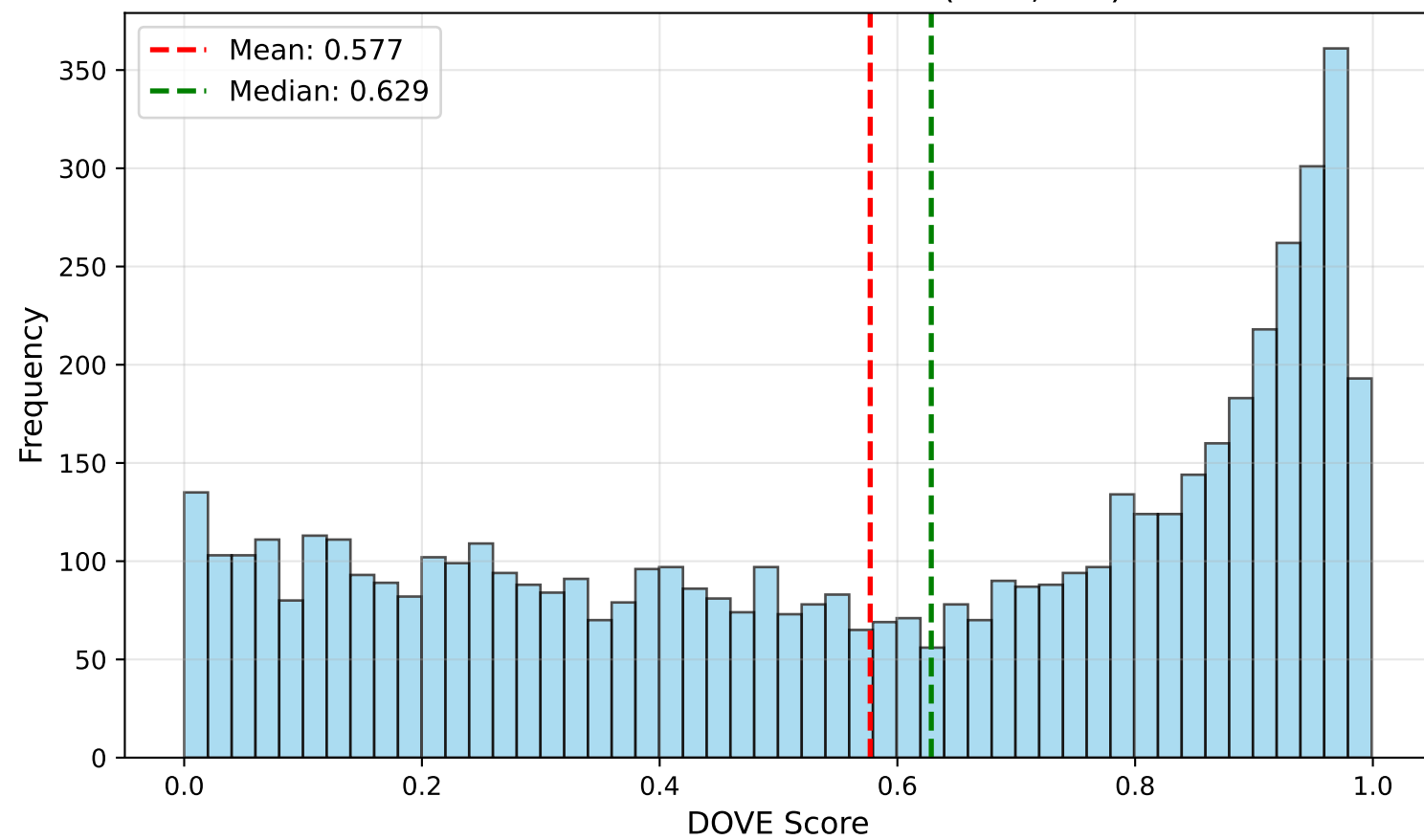
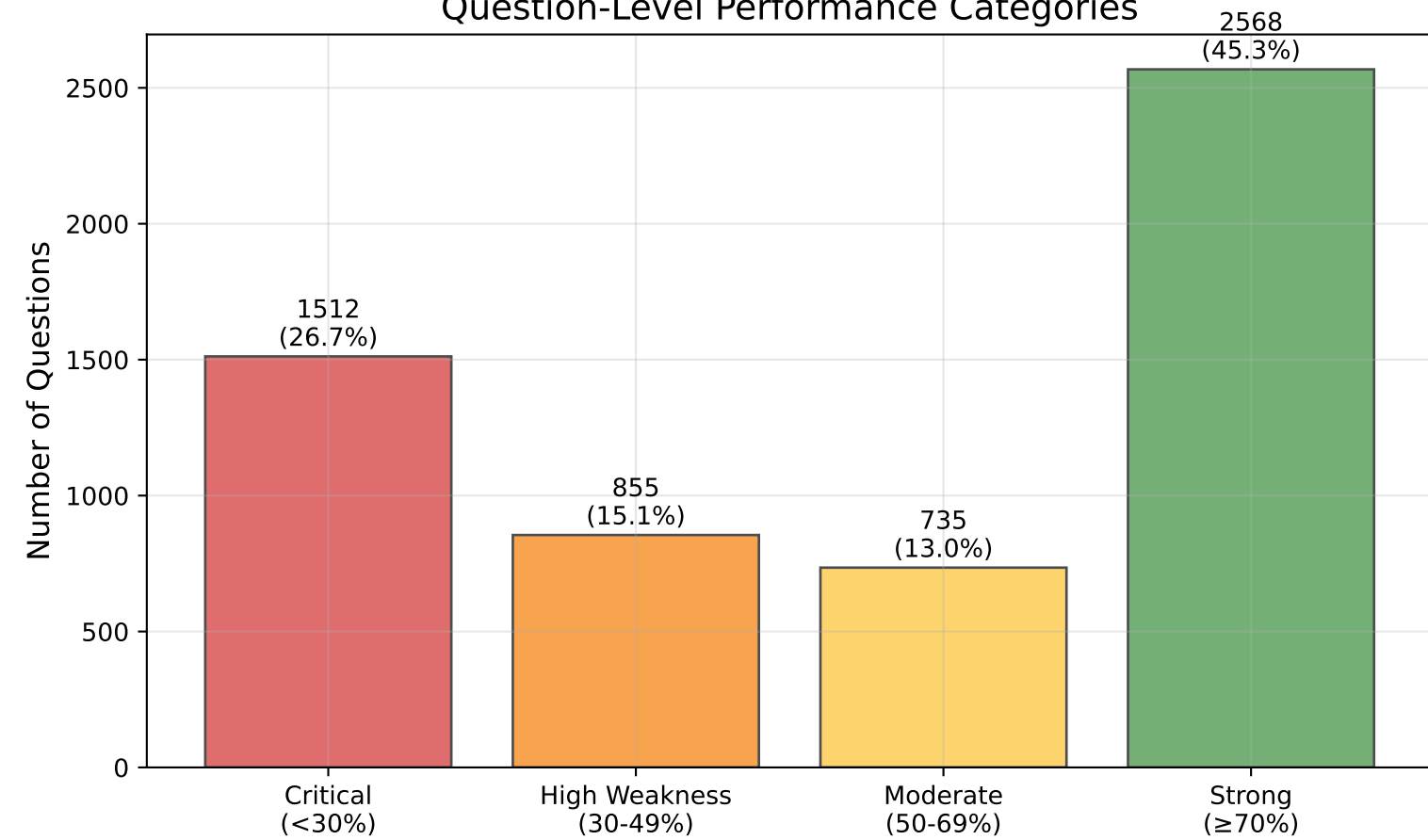


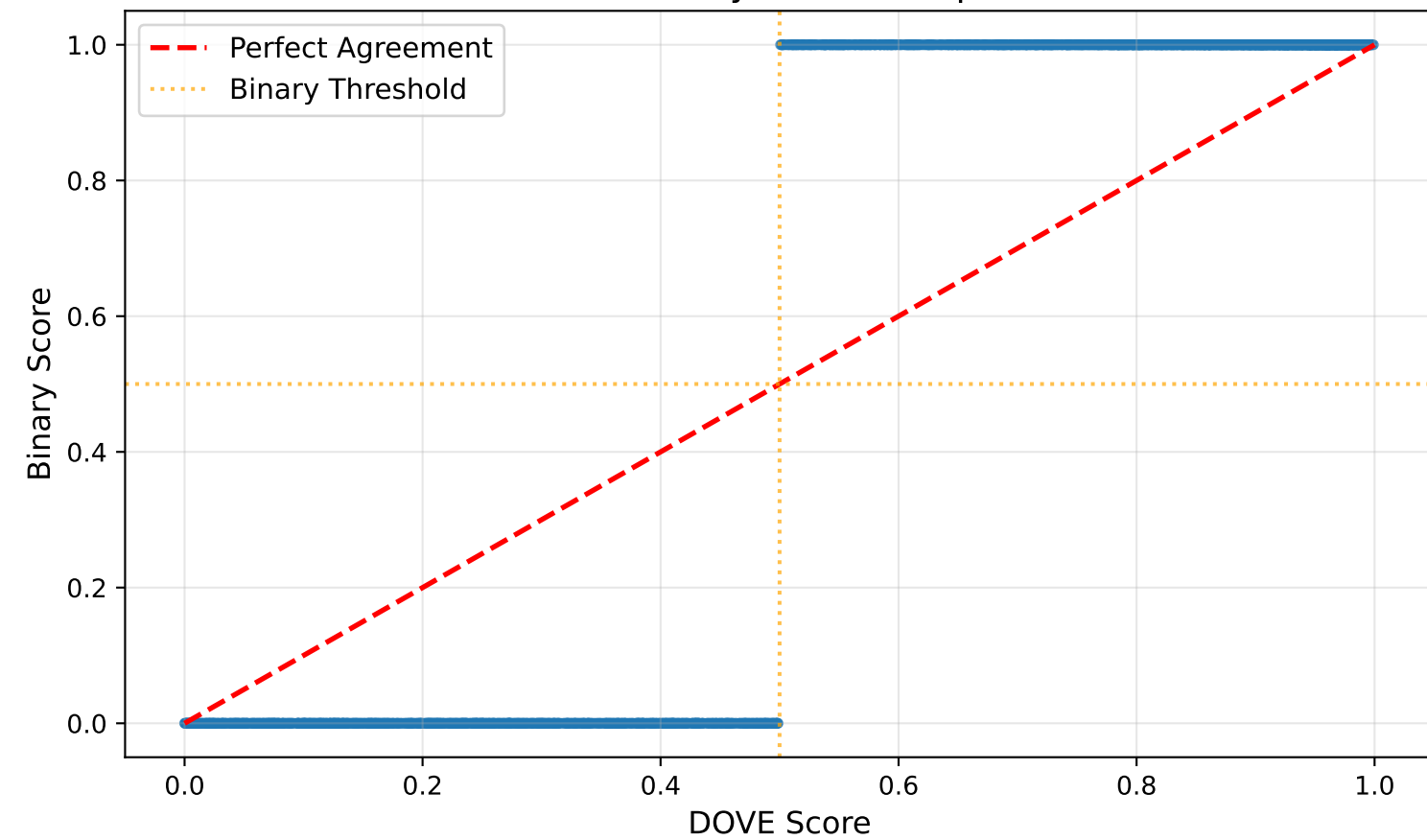
Distribution of DOVE Scores (n=5,670)



Question-Level Performance Categories



DOVE vs Binary Score Comparison



Score Distribution Comparison

