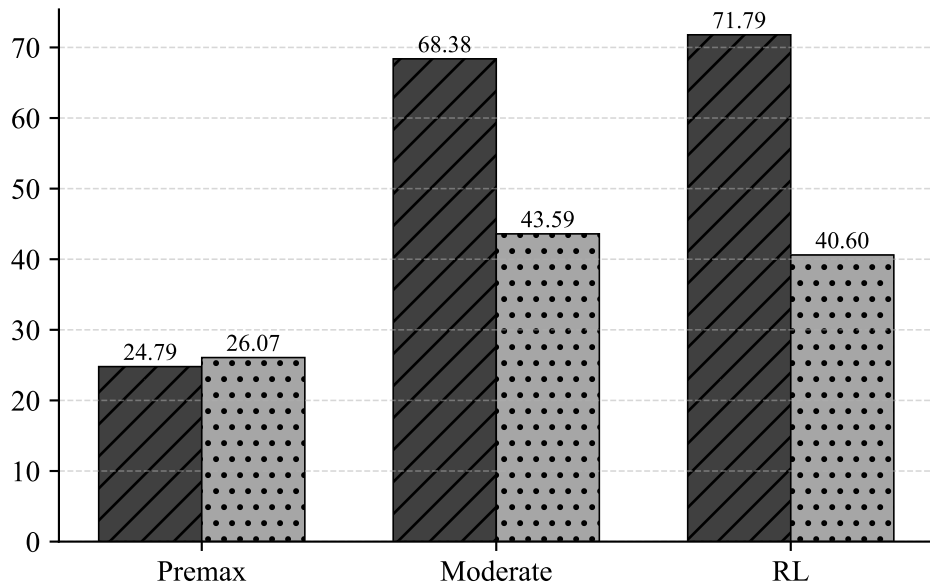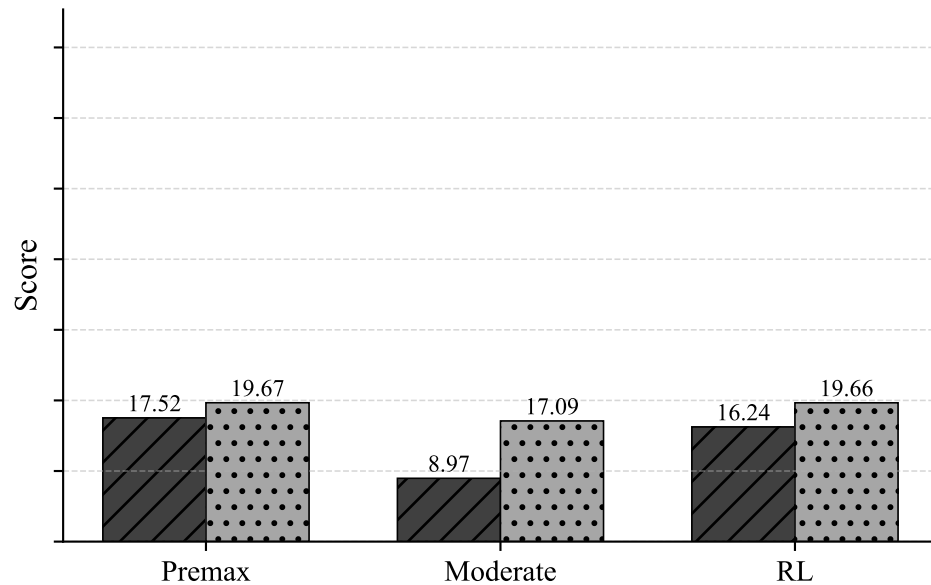Task 1 Performance by Stage (ID vs OOD)