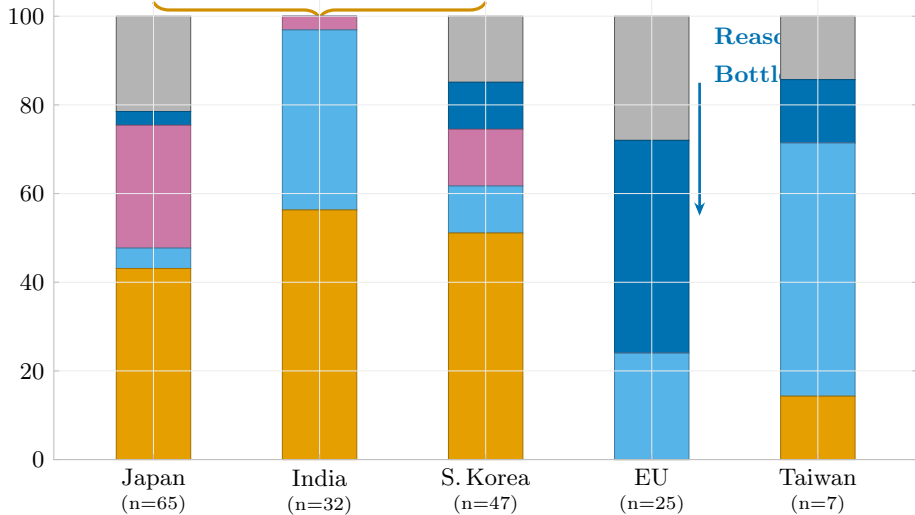


Proportion of Failures (%)

Perception Bottleneck

Reasoning Bottleneck



Non-Latin Script

OCR/Text

Math/Symbol

Reasoning

Others