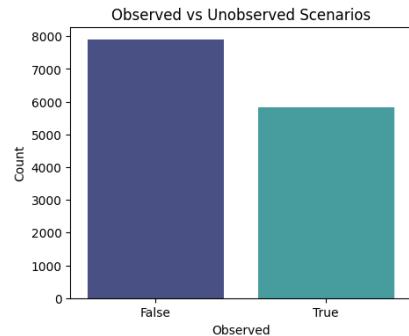


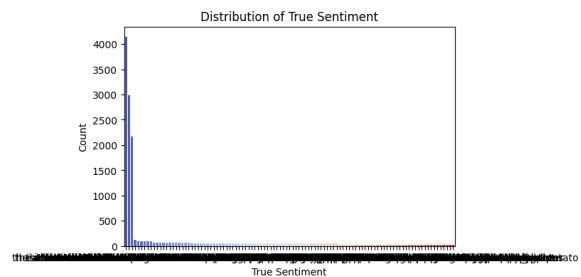
Theory of Mind Dataset Analysis — Graph Explanations

Distribution of True Sentiment



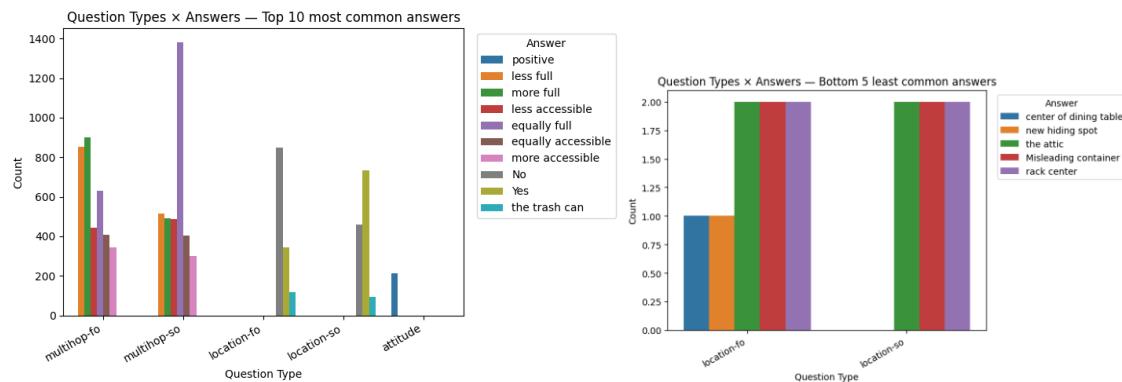
This graph shows how different sentiment labels (like 'positive', 'negative', or 'neutral') are distributed across the dataset. Most entries are strongly negative, suggesting the dataset contains many situations involving negative emotions or evaluations.

Observed vs Unobserved Scenarios



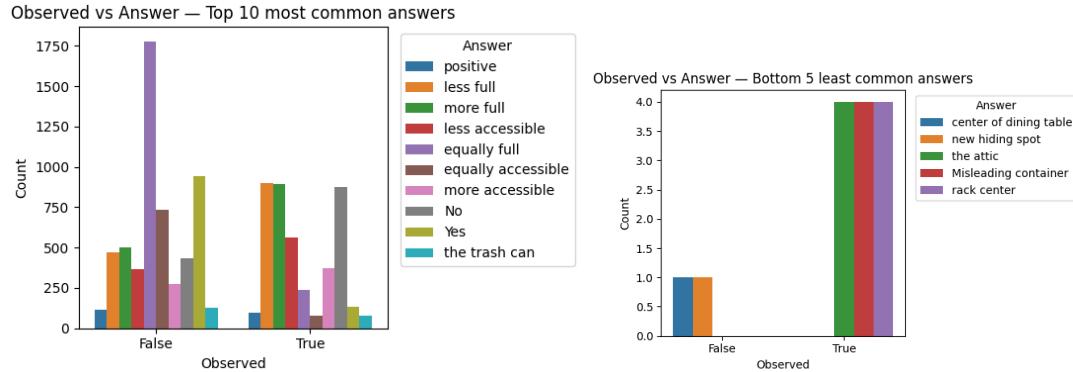
This chart compares how many scenarios were observed (the observer saw the action) versus unobserved (the observer did not see the action). There are slightly more unobserved cases, meaning false-belief reasoning may be common in the dataset.

Question Types x Answers — Top 10 most/ 5 least common answers



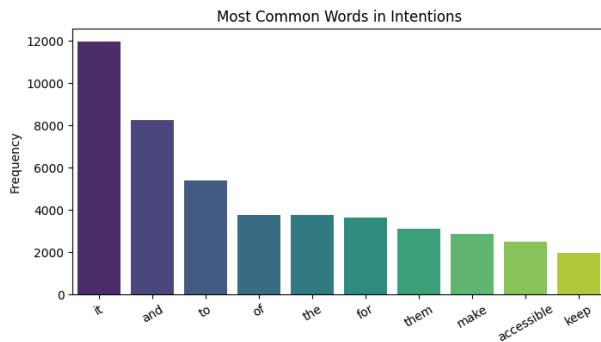
The first plot displays the most frequent answer types for each question category (e.g., attitude, location, multihop). It highlights patterns in how the model or participants answered standard ToM questions. The second is the same but for 5 least frequent answers.

Observed vs Answer — Top 10 most common answers



The first chart compares the top 10 most common answers between observed and unobserved scenarios. It shows how witnessing an action affects judgments or interpretations — key to understanding perspective-taking. The second is the same but for 5 least frequent answers.

Most Common Words in Intentions



This bar chart highlights the most frequent words used in intention descriptions. Common words like 'it', 'make', and 'accessible' suggest the dataset often describes purposeful, goal-driven actions.