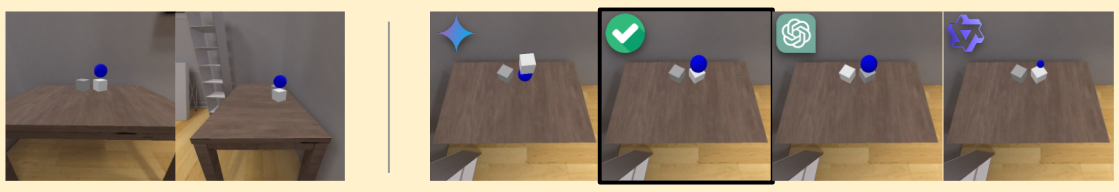
Spatial

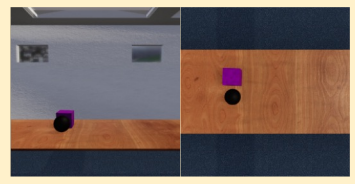
Spatial Positioning (SP)

Q: Given the front view and side view of the two objects, select the correct top view.



Spatial Relation (SR)

Q: Where is the black sphere relative to the purple cube from the front view?



GPT-4o: behind
Gemini 1.5: left
Qwen-VL: bottom
Answer: front

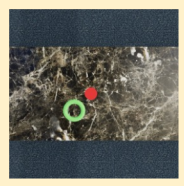
Spatial Occupancy (SE-O)

Q : Which object appears to be larger in size?

GPT-4o: sphere
Gemini 1.5: cube
Qwen-VL: cone
Answer: cylinder

Spatial Vacancy (SE-V)

Q : Can the cylinder be placed into the ring from the top?



GPT-4o: No
Gemini 1.5: No
Qwen-VL: No
Answer: Yes

Quantity

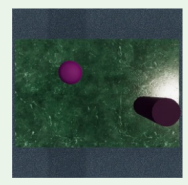
Discrete Quantity (DQ)

Q: How many orange objects are there?

GPT-4o: 5
Gemini 1.5: 5
Qwen-VL: 5
Answer: 6

Continuous Quantity (CQ)

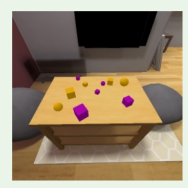
Q: Which object has a deeper color?



GPT-4o: cylinder
Gemini 1.5: sphere
Qwen-VL: cylinder
Answer: cylinder

Relative Quantity (RQ)

Q: Which color has more objects?

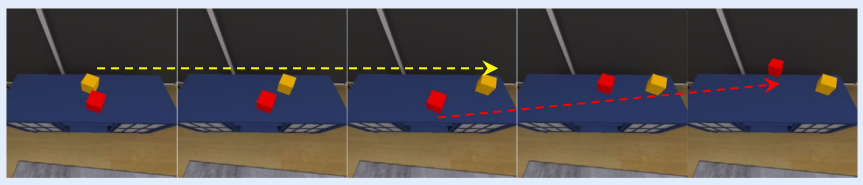


GPT-4o: purple
Gemini 1.5: purple
Qwen-VL: purple
Answer: same

Temporal

Temporal Positioning (TP)

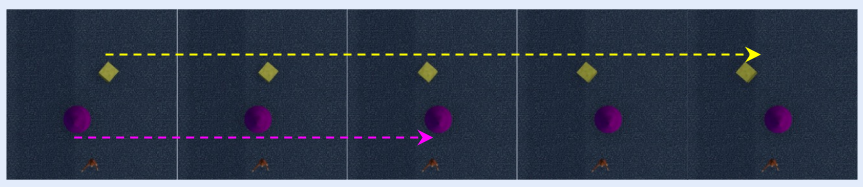
Q: Which object starts moving first based on the trajectory?



GPT-4o: same time
Gemini 1.5: red cube
Qwen-VL: red cube
Answer: yellow cube

Temporal Extension (TE)

Q: Which object appears to move for a longer duration?

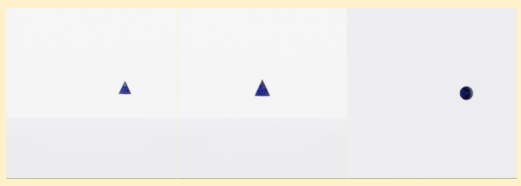


GPT-4o: same duration
Gemini 1.5: purple cone
Qwen-VL: purple cone
Answer: yellow cube

Visual

Shape (V-S)

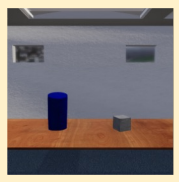
Q: What is the shape of the object?



GPT-4o: cylinder
Gemini 1.5: sphere
Qwen-VL: sphere
Answer: cone

Material (V-M)

Q: Do they share the same material?



GPT-4o: No
Gemini 1.5: No
Qwen-VL No
Answer: Yes

Color (V-C)

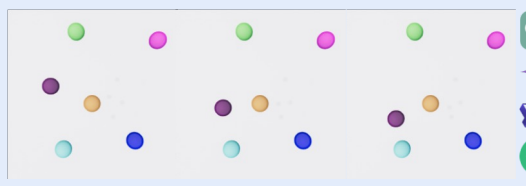
Q: What color is the object?

GPT-4o: white
Gemini 1.5: white
Qwen-VL white
Answer: gray

Motion

Motion Identification (MI)

Q: Which object is moving?



GPT-4o: blue
Gemini 1.5: orange
Qwen-VL: blue
Answer: purple

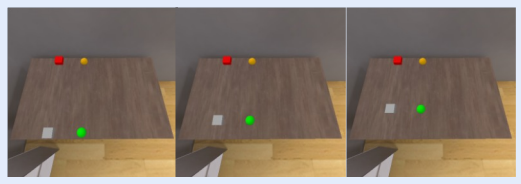
Motion Direction (MD)

Q: Towards which object is the cyan sphere moving?

GPT-4o: orange
Gemini 1.5: none
Qwen-VL: orange
Answer: yellow

Motion Speed (MS)

Q: Which object is moving faster?



GPT-4o: green
Gemini 1.5: green
Qwen-VL: gray
Answer: same

Motion Trajectory (MT)

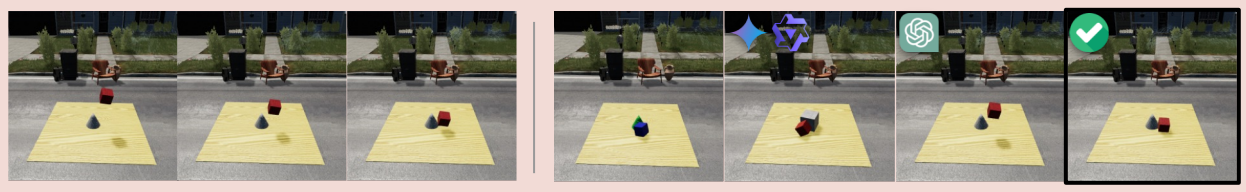
Q: What is the moving trajectory of the cyan sphere?

GPT-4o: circle
Gemini 1.5: line
Qwen-VL: line
Answer: Square

Mechanistic

Mechanistic Simulation, Intuitive Physics (M-IP)

Q: Given a sequence of images showing consecutive states of the environment, which of the following images is most likely to be the next state after?



Mechanistic Simulation, Agent Navigation (M-Nav)

Q: Given the start state, which is most likely to be the final state after the robot **moves backward**?

Mechanistic Simulation, Agent Manipulation (M-Man)

Q: Given the start state, which is most likely to be the final state after the robot arm **pushes the object**?

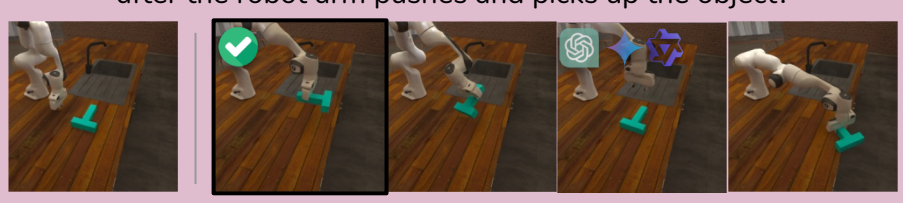
Transitive

Transitivity, Agent Navigation (T-Nav)

Q: Given the start state, which is most likely to be the final state after the vehicle **turns left and moves forward**?

Transitivity, Agent Manipulation (T-Man)

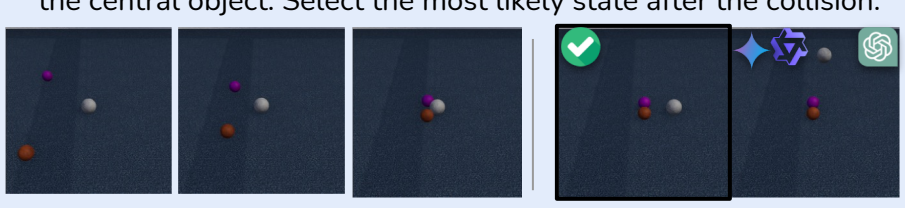
Q: Given the start state, which is most likely to be the final state after the robot arm pushes and picks up the object?



Compositional

Compositionality, Multi-Object Intuitive Physics (C-IP)

Q: Two spheres with equal mass and same velocity move towards the central object. Select the most likely state after the collision.



Compositionality, Multi-Agent Manipulation (C-Man)

Q: Given the start state, predict the final state after both robotic arms successfully and simultaneously lift the object.

Perception
WM-ABench
Prediction