

Category 1: Numerical

Question 1 (1 point): How many pillows are on the bed?



Response: Print 4 in terminal.

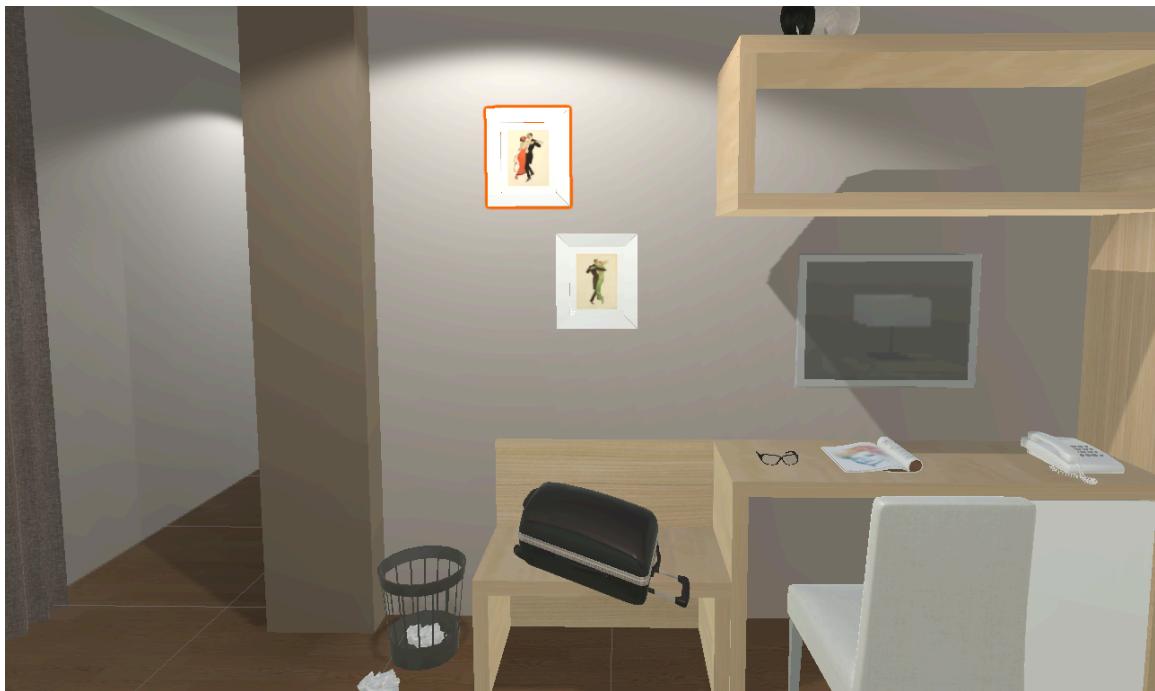
Category 2: Object Reference

Question 2 (1 point): Find the bedside table farthest from the window.



Response: Bounding box of the object

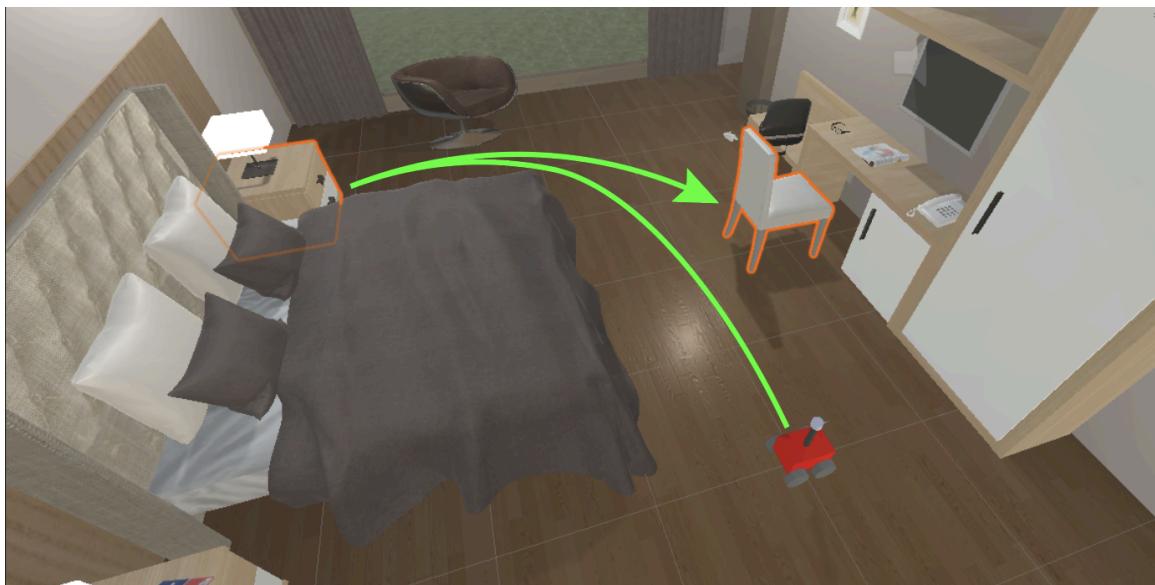
Question 3 (1 point): Find the picture above the suitcase furthest from the floor.



Response: Bounding box of the object

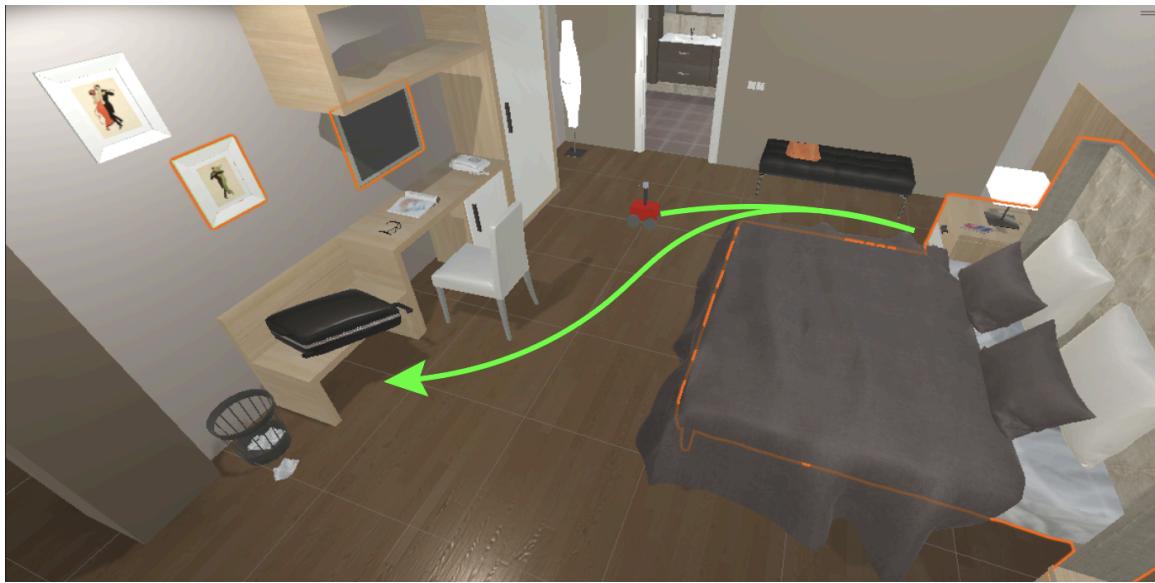
Category 3: Instruction Following

Question 4 (2 points): Go to the bedside table closest to the window and stop at the chair closest to the TV.



Response: trajectory (series of waypoints)

Question 5 (3 points): First, go near the bedside table closest to the bench, then take the path between the TV and the bed to the picture closest to the TV.



Response: trajectory (series of waypoints)

Note: The number of points of a question is based on the number of steps in the response. For example, Question 5 expects 3 steps in the response and is worth 3 points. The score is calculated as the number of correctly responded steps.