GPT-4o

Q: How many balls are there in this image?

A: There are 10 balls. ✗

Q: The guest will sit on the right side of the table, near the yellow cup. Where should we serve the apple?

A: Apple is at [445,535], and should be placed at [450,150]. ✗

GPT-4o with Autonomous Imagination

Q: How many balls are there in this image?

Creating the imagination space.
Controlling cursor.

Reasoning in Imagination Space
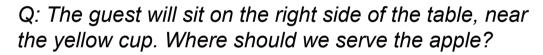
I should count by ignoring balls.
Let's start from the red and white one.

Reasoning in Imagination Space
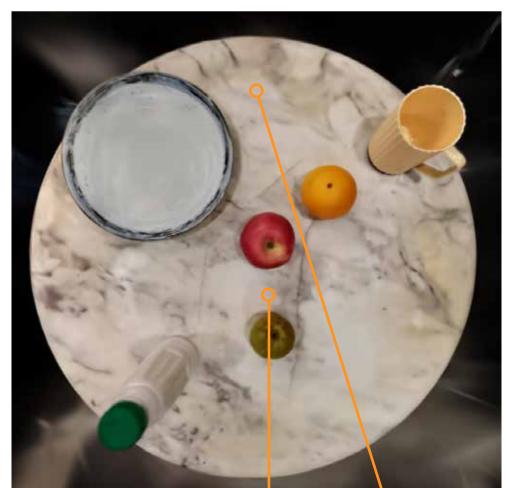
I have ignored the red and white ball.
Let's count the white ball on the left.

Reasoning in Imagination Space

A: There are 11 balls. ✓

Q: The guest will sit on the right side of the table, near the yellow cup. Where should we serve the apple?

Creating the imagination space.
The apple is right at the cursor.

Reasoning in Imagination Space

I have located the apple at [450,450], ✓ *moving it to the right side.*

Reasoning in Imagination Space

This location looks better, but I will try to find a better location.

Reasoning in Imagination Space

A: I think this is a nice spot.
Target location: [705,376].