

Level 3 - Coupled Color-Count

Question: You are presented with a visual reasoning puzzle consisting of a 3×3 grid of images, where one image is missing in one of the nine positions.

Below the grid are four candidate options labeled A, B, C, and D, arranged horizontally from left to right.

Exactly one option correctly completes the grid.

Task: 1. The grid follows a uniform logical rule, which governs the relationships between the images across the entire grid; 2. In this puzzle, focus on the following attributes: number of elements (count) and color, which may vary independently or in relation to one another; 3. Carefully analyze how the attributes vary in each image and across the grid to infer the common rule that governs them; 4. Choose the option that best follows this rule and completes the grid in a logically consistent way.

In your reply, follow this format exactly:

Reasoning: Describe the main rule or pattern and explain why your answer fits it.

Answer: [A, B, C, or D]



Ground Truth: D ✓

GPT-5-Thinking

Reasoning: Count the number of items in each picture.

Row 1 totals $3 + 4 + 1 = 8$ items.

Row 2 totals $6 + 2 + 1 = 9$ items.

Thus, the rule is that the total number of objects per row increases by 1 each row, so Row 3 must total 10. The existing images in Row 3 have 3 and 2 items, so the missing cell must contain 5 items ($10 - (3 + 2) = 5$). Among the options, only D shows five objects, and it also matches the warm-color tendency of the first column.

Wrong reasoning with correct answer

Model Answer: D ✓

Coupled rule failure: the model ignores the count–color coupling and instead uses a row-sum heuristic (“8, 9, 10”) plus a vague warm-color bias, arriving at the correct option D for the wrong reasons.