

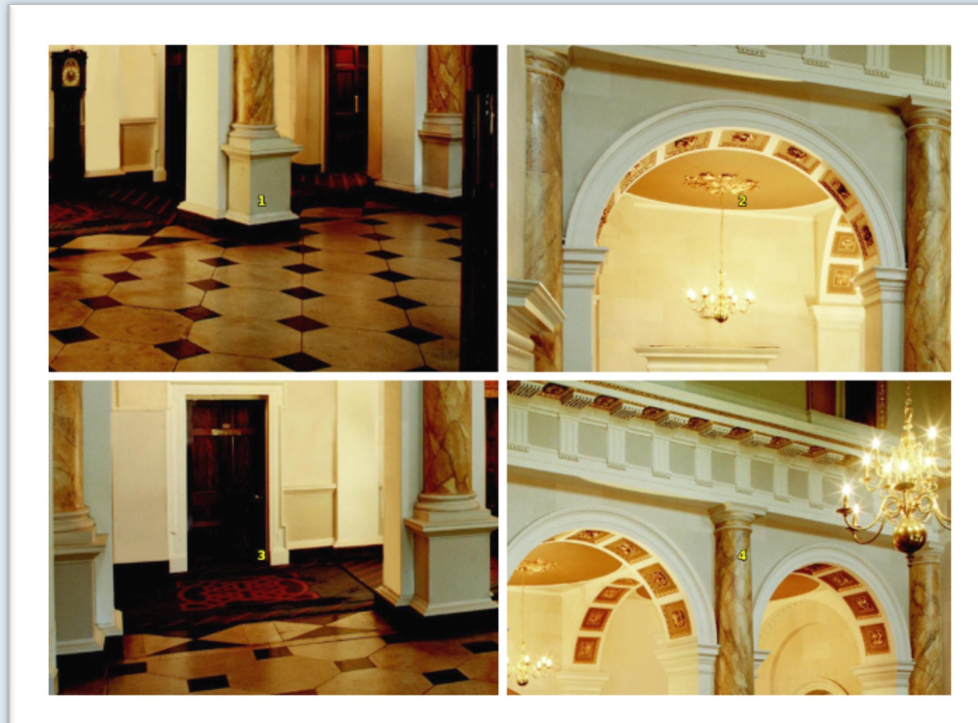


ThinkMorph Multimodal Reasoning

Jigsaw Assembly

The image below is a 2x2 jigsaw puzzle...
Your task is to determine the correct arrangement ...

- A. ...
- B. ...
- C. ...
- D. ...



<think>... analyze the visual content of each piece and determine their correct positions based on natural image flow ... </think>



<think>... There are **no visible discontinuities** ... harmoniously unified... </think>
<answer> C</answer>

Spatial Navigation

Your goal is to guide a player from the start to the goal on a grid map while avoiding holes ...
Provide your solution as a sequence of moves ...



<think>Let's analyze the maze step by step: Step 1: Identify the grid and obstacles The player starts at the top-left corner (let's call this position (1,1))... </think>

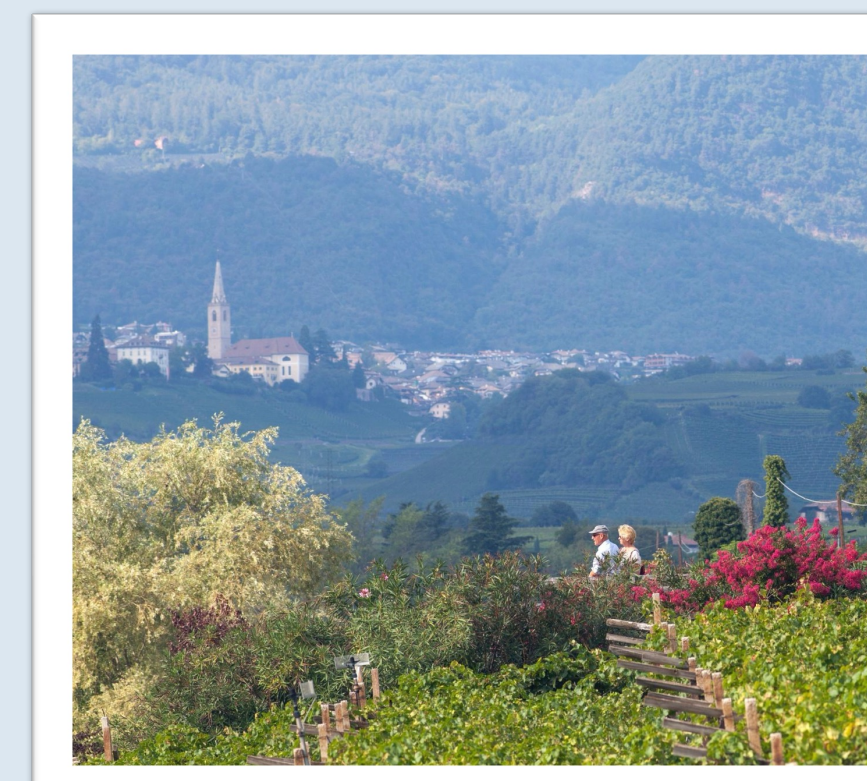


<think>...carefully verify the path I plotted. **The red arrows** start at the top-left corner ... </think> <answer> \boxed{D,D,D,R,R,R} </answer>

Visual Search

What is the color of the man's cap?

- A. green B. white
- C. orange D. blue



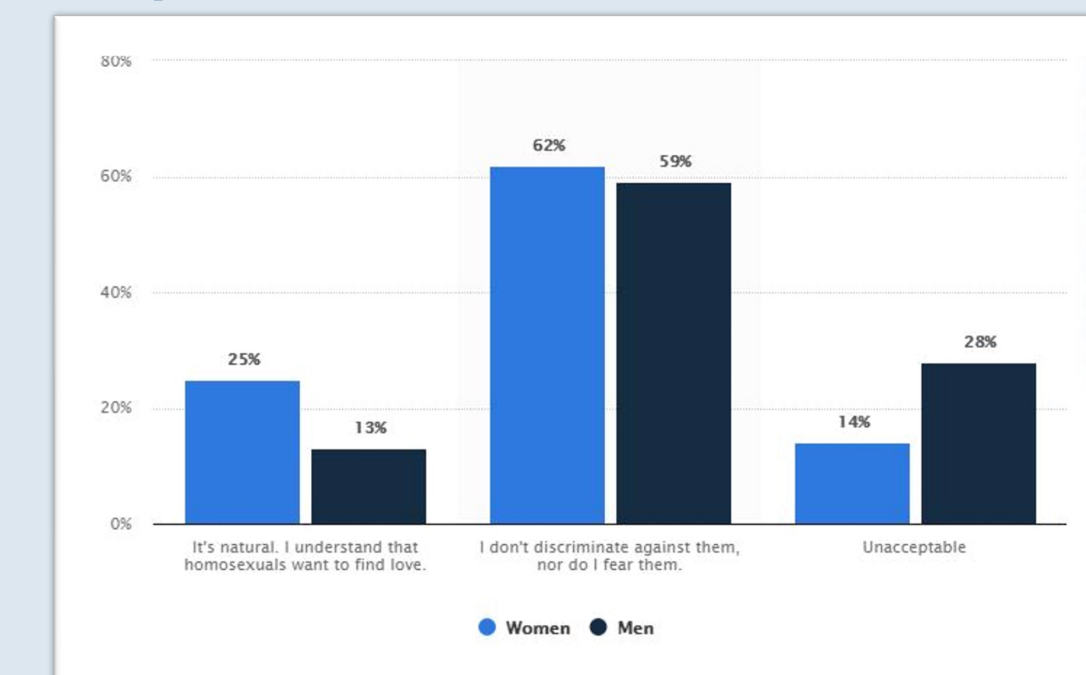
<think>.....The **red bounding box** should be placed around the **man's head** and cap....</think>



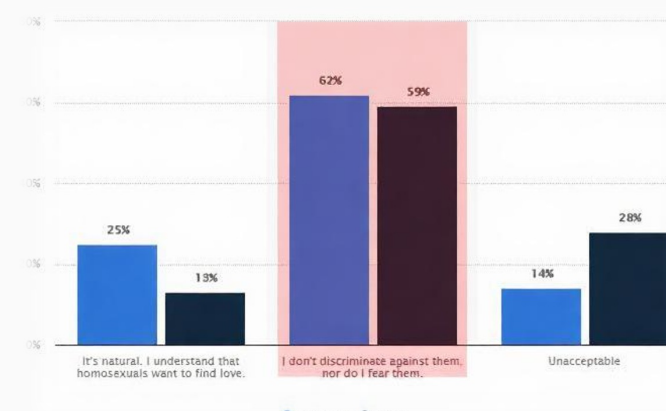
<think>... now highlighting the man ... focusing on the highlighted area, ... **the man's cap is gray**, so the correct answer is C. </think> <answer>C</answer>

Chart Refocus

Which attitude represent the smallest gap between women and men?



<think>To answer ... I **focus on** these two specific bars and labels ... </think>



<think> ... I analyze the highlighted area in the **visualization** that **focuses on the second category: 'I don't discriminate against ...'** </think> <answer>I don't discriminate against them, nor do I fear them.</answer>

Emergent Properties

1 Unseen Visual Manipulations



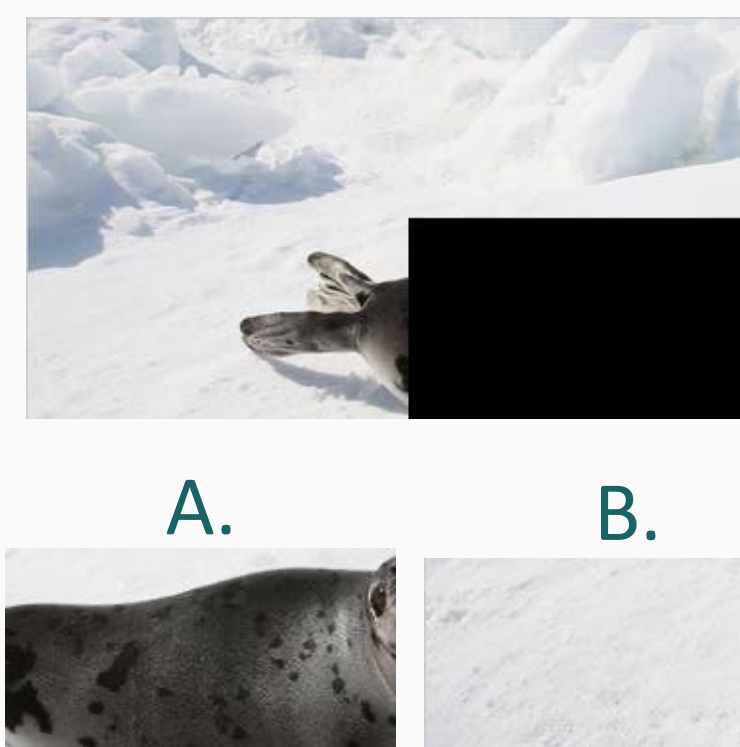
Is the bell pepper in the picture red or yellow?
A: Red B: Yellow
(From MMVP)

<think>To determine whether the bell is red or yellow, I need to **closely examine the color** of the pepper shown ...



<think>... I closely examine the image provided, which features a close-up shot as its central focus... Based on this detailed visual inspection, I can confidently conclude that the bell pepper is **yellow**... </think>
<answer>B</answer>

Zoom-in



Given the first image with the lower right corner missing, ... compare the edges of the images. Select from the following choice.
(From BLINK)

<think> ... To determine the correct assembly, I focus on the continuity between the two pieces ... **reconstructing** the seal as it would appear in reality.</think>



<think>... I can examine the result for visual coherence ... overall composition **looks natural** ... as in a single, unaltered photo....</think> <answer>A</answer>

Image Inpainting

2 Autonomous Mode Switching



Interleaved Thought

Text thought

OR

e.g., 5.3% Switch
7.29% Gain

3 Superior Test-Time via Diversified Thoughts

