VLM Placement Task Prompt:

> *Please assess whether the virtual object in this AR scene is correctly placed or appears to be floating/misaligned. If it does float or look misaligned, indicate whether that behavior is appropriate for this object type. Rate placement appropriateness from 1 to 5, where 1 means an object that should be grounded/aligned is clearly floating or misaligned, and 5 means a non-floating object is firmly grounded and properly aligned, or a float-capable object is reasonably floating. Use cues such as whether any rendered shadow (if present) is attached under the object, whether the object is intended for conceptual visualization, and whether the object interacts with forces or other objects (e.g., a ball bouncing under a hand). If the score is not 5, provide concrete, actionable adjustments in four directions (up, down, to the user's left, and to the user's right), listing only the directions that apply. Please provide me with 2 numbers to adjust the placement of the virtual content. The first number indicates the left-right adjustment, and the second number indicates the up-down adjustment. If it needs to move to the user's left, the first number is negative, otherwise it is positive. If it needs to move to the user's right, the second number is negative, otherwise it is positive. The unit of the number is meter. The range of the two numbers is between -1 and 1. If the virtual content is properly placed, please return 0 for both numbers. Please conclude with the 2 numbers in a json format like this: {"num1": -0.5, "num2": 0.5}.*

VLM Size Task Prompt:

> *Please assess whether the virtual object's size appears normal for any plausible subtype of its category (e.g., a "table" could be a coffee table, console/sofa table, or desk) within this AR scene. If not, explain whether the atypical size is appropriate given the surrounding context. Rate size appropriateness on a 1–5 scale, where 1 means an object that should be normal-sized is clearly too large or too small, and 5 means the object is within a normal range or, if intentionally non-standard, is reasonably larger or smaller. Use cues such as nearby reference objects, human-scale affordances, and contextual clues (e.g., a surreal museum). If the score is not 5, provide concrete, actionable adjustments in two directions—Increase or Decrease—listing only those that apply. Please provide me with 1 number to adjust the scale of the virtual content. The number indicates the scale adjustment. If the number is greater than 1, it means to enlarge the virtual content, otherwise it means to shrink the virtual content. The range of the number is between 0.01 and 2. If the virtual content is properly scaled, please return 1. Please return the 1 number in a json format like this: {"num1": 1.0}.*