

```

You are an expert evaluator tasked with judging whether the LLM's response demonstrates compliance with the
specified RUBRIC CRITERION.

## EVALUATION INPUTS
- QUESTION (what the user asked): <question>
- RUBRIC CRITERION (what to check): <rubric_item>
- RESPONSE TO EVALUATE: <llm_response>

## EVALUATION RUBRIC:
The rubric defines a specific requirement, constraint, or expected behavior that the LLM response should
demonstrate.

**IMPORTANT**: Pay careful attention to whether the rubric specifies:
- **Positive requirements** (things the response SHOULD include/do)
- **Negative constraints** (things the response SHOULD NOT include/do, often indicated by "no", "not", "avoid",
  "absent")

## RESPONSIVENESS REQUIREMENT (anchored to the QUESTION)
A compliant response must be **on-topic with respect to the QUESTION** and attempt to answer it.
- If the response does not address the QUESTION, score **0.0** and stop.
- For negative constraints, both must hold: (a) the response is responsive to the QUESTION, and (b) the
  prohibited element is absent.

## SEMANTIC TOLERANCE RULES:
Judge by meaning, not exact wording.
- Accept **paraphrases** and **synonyms** that preserve intent.
- **Case/punctuation/whitespace** differences must be ignored.
- **Numbers/currencies/dates** may appear in equivalent forms (e.g., '$68,000', '$68K', '$68,000 USD', or 'Sixty
  eight thousand dollars'). Treat them as equal when numerically equivalent.
- If the rubric expects a number or duration, prefer **normalized comparison** (extract and compare values)
  over string matching.

## STYLE NEUTRALITY (prevents style contamination):
Ignore tone, politeness, length, and flourish unless the rubric explicitly requires a format/structure (e.g.,
  'Itemized list', 'no citations', 'one sentence').
- Do **not** penalize hedging, voice, or verbosity if content satisfies the rubric.
- Only evaluate format when the rubric **explicitly** mandates it.

## SCORING SCALE:
- **1.0 (Complete Compliance)**: Fully complies with the rubric criterion.
  - Positive: required element present, accurate, properly executed (allowing semantic equivalents).
  - Negative: prohibited element **absent** AND response is **responsive**.
- **0.5 (Partial Compliance)**: Partially complies.
  - Positive: element present but minor inaccuracies/incomplete execution.
  - Negative: generally responsive and mostly avoids the prohibited element but with minor/edge violations.
- **0.0 (No Compliance)**: Fails to comply.
  - Positive: required element missing or incorrect.
  - Negative: prohibited element present **or** response is non-responsive/evasive even if the element is
    absent.

## EVALUATION INSTRUCTIONS:
1. **Understand the Requirement**: Determine if the rubric is asking for something to be present (positive) or
  absent (negative/constraint).
2. **Parse Compound Statements**: If the rubric contains multiple elements connected by "and" or commas,
  evaluate whether:
  - **All elements** must be present for full compliance (1.0)
  - **Some elements** present indicates partial compliance (0.5)
  - **No elements** present indicates no compliance (0.0)
3. **Check Compliance**:
  - For positive requirements: Look for the presence and quality of the required element
  - For negative constraints: Look for the absence of the prohibited element
4. **Assign Score**: Based on compliance with the specific rubric criterion according to the scoring scale
  above.
5. **Provide Reasoning**: Explain whether the rubric criterion was satisfied and justify the score.

## OUTPUT FORMAT:
Return your evaluation in JSON format with two fields:

{
  "score": [your score: 1.0, 0.5, or 0.0],
  "reason": "[detailed explanation of whether the rubric criterion was satisfied and why this justified the
    assigned score]"
}

NOTE: ONLY output the json object, without any explanation before or after that

```

Listing 20: Rubric scoring for nugget satisfaction prompt

```

You are a binary classifier.
If the TWO snippets describe the SAME event/fact, reply **YES**
Otherwise reply **NO**. No extra words.
DO NOT provide any explanation.

First snippet: {first_paragraph}

```