

- LLM response should mention: <event N>

Example JSON:

```
{
  "question": "How did my focus on different aspects of property investment and management develop throughout our conversations in order? Mention ONLY and ONLY ten items.",
  "answer": "Your focus on property investment and management developed in this sequence: 1) Initial engagement with the local agent and preparation for property viewings, 2) Evaluation of property financials including ROI and rental income potential, 3) Exploration of financing options and mortgage concerns, 4) Handling contractor performance and repair negotiations, 5) Understanding zoning regulations and permit requirements for property conversions, 6) Planning and prioritizing renovations and investment risks for multi-family properties, 7) Detailed preparation for property viewings involving both agent and contractor, 8) Negotiation strategies and cash flow implications related to repair costs, 9) Final renovation project completion steps and portfolio diversification strategies, 10) Consideration of commercial property types and location factors for long-term investment.",
  "ordering_tested": [
    "1st: Agent interaction and viewing preparation",
    "2nd: Property financial evaluation",
    "3rd: Financing and mortgage concerns",
    "4th: Contractor management",
    "5th: Zoning and permits",
    "6th: Renovation planning and investment risks",
    "7th: Viewing preparation with agent and contractor",
    "8th: Repair cost negotiation and cash flow",
    "9th: Renovation completion and portfolio diversification",
    "10th: Commercial property and location considerations"
  ],
  "source_chat_ids": [],
  "rubric": [
    "LLM response should mention: Agent interaction and viewing preparation",
    "LLM response should mention: Property financial evaluation",
    "LLM response should mention: Financing and mortgage concerns",
    "LLM response should mention: Contractor management",
    "LLM response should mention: Zoning and permits",
    "LLM response should mention: Renovation planning and investment risks",
    "LLM response should mention: Viewing preparation with agent and contractor",
    "LLM response should mention: Repair cost negotiation and cash flow",
    "LLM response should mention: Renovation completion and portfolio diversification",
    "LLM response should mention: Commercial property and location considerations",
    "Presents the events in the correct chronological order"
  ]
}
```

4. Information Extraction

Objective: LLM should answer the questioned facts correctly.

Rubric pattern:

- Instantiate one criterion per fact directly from the ideal answer, using the stem “LLM response should state/mention:”

Example JSON:

```
{
  "question": "What amount did I say I had set aside as my initial capital for investing?",
  "ideal_answer": "You said you had set aside $50,000 as your initial capital.",
  "source_chat_ids": [],
  "rubric": [
    "LLM response should state: $50,000"
  ]
}
```

5. Instruction Following