

---

```

Return your analysis in this exact JSON format:
[{"capability": "preference_following", "batch_numbers": 1,"bullet_numbers": 17,
    "bullet_points": "***Preference Statement:** I prefer materials that balance cost and performance;
        choose 3.5 lb/ft$^2$ MLV despite 20% higher price."}
]

Important formatting notes:
- The "batch_numbers" and "bullet_numbers" correspond to each other positionally
- "1" and "17" means: Batch 1 Bullet 17

Select ONLY <bullet_number> bullet points total that would generate the highest quality preference following
questions.

NOTE: Only output the list without any explanation before or after the list.

```

**Listing 5:** Candidate selection preference following prompt

```

This is a plan that contains detailed bullet points about a topic. This plan is used to generate realistic
chat conversations between a user and an AI assistant, which are then used to evaluate the long-term
memory capabilities of LLMs.
Your task is to analyze this plan and select GROUPS of related bullet points that would be most effective for
testing event ordering abilities when incorporated into chat conversations.
Event ordering tests whether the LLM can recall the chronological order in which events or topics were
MENTIONED in the conversation, regardless of when the actual events occurred in real life.
Analyze this plan and identify GROUPS of 8-12 or more related bullet points that represent the same topic/
theme mentioned across different batches, enabling testing of mention-order recall and conversation
sequence understanding.

## INPUT DATA
- **PLAN**: <plan>

## CRITICAL REQUIREMENT: EARLY BATCH PRIORITIZATION
**SELECTION PRIORITY ORDER:**
1. **Groups starting in Batches 1-3 (HIGHEST PRIORITY)**: Select 70-80% of your groups with first mention in
   early batches
2. **Groups spanning early to middle batches (MEDIUM PRIORITY)**: Select 10-20% of groups that bridge early-to
   -middle timeline
3. **Groups from later batches only (LOW PRIORITY)**: Select only 5-10% from purely later batches

Focus on bullet point groups that show:
- **Same person mentioned multiple times**: Different interactions or mentions of the same person across
   batches
- **Same component/process discussed repeatedly**: Multiple mentions of the same equipment, material, or
   process
- **Same location/venue referenced**: Multiple mentions of the same place or address
- **Same decision/topic revisited**: The same subject brought up in different conversation sessions
- **Same problem/solution mentioned**: Multiple references to the same issue across different times
- **Same financial item tracked**: Multiple mentions of the same cost, budget item, or expense

Prioritize bullet point groups that:
- Contain 8-12 mentions of the same topic across different batches
- Enable questions about "In what order events X,Y,Z,... happen?" Or ...
- Allow testing of conversation chronology rather than real-world event chronology
- Test recall of mention sequence: "Which did I talk about first, second, third?"
- Focus on the order topics appeared in conversation, not when events actually happened
- Create opportunities to test conversational memory rather than factual timeline memory

Return your analysis in this exact JSON format where each object contains 3+ bullet points about the same
topic across different batches:
[{"capability": "event_ordering", "batch_numbers": "1, 3, 5, 7", "bullet_numbers": "22, 18, 11, 27",
    "bullet_points": ". **Character & Relationship:Acoustic Consultant:** Met Rami Al-Hassan at Bahrain
        Acoustic Expo on Feb 20; he recommended HVAC silencing at $2,200. | . **Character & Relationship
        :Acoustic Consultant:** Rami conducted mid-project site visit April 3, advised on bass trap
        repositioning to improve 5 dB absorption. | . **Character & Relationship:Acoustic Consultant:** 
        Rami praised progress in May 1 email; suggested minor EQ tweaks. | . **Character & Relationship:
        Acoustic Consultant:** Rami praised final results August 20; recommended ongoing maintenance and
        periodic EQ checks."}
]

Important formatting notes:
- The "batch_numbers" and "bullet_numbers" correspond to each other positionally
- "1, 2, 3, 5, 7" and "22, 18, 11, 27" means: Batch 1 Bullet 22, Batch 3 Bullet 18, Batch 5 Bullet 11, Batch 7
  Bullet 27
- Each object must contain 8-12 related bullet points separated by " | "
- Use comma-separated values for batch_number and bullet_number
- All bullet points must reference the same topic/person/component/theme
- Focus on groups that enable mention-order questions
- Test conversational chronology, not real-world event chronology

Select 8-10 groups of bullet points that would enable the most sophisticated mention-order and conversation
sequence questions.

CRITICAL NOTE: DO NOT consider bulletpoint names for selecting the bullets. ONLY consider the bullets contents
.

NOTE: Only output the list without any explanation before or after the list.

```

**Listing 6:** Candidate selection event ordering prompt